DOCUMENT RESUME

ED 024 634 24 SP 001 886

By- Turner, Richard L.
Differential Association of Elementary School Teacher Characteristics with School System Types. Final Report.
Indiana Univ., Bloomington.
Spons Agency- Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No- BR-5-0345
Pub Date Sep 68
Contract- OEC-5-10-059
Note- 154p.
EDRS Price MF-$0.75 HC-$7.80
Descriptors- Academic Achievement, Behavior Rating Scales, Community Characteristics, Economic Factors, *Educational Research, *Elementary School Teachers, Individual Characteristics, Models, Organization, Problem Solving, School Districts, School Surveys, *School Systems, Social Characteristics, *Socioeconomic Status, Task Analysis, *Teacher Characteristics, Teaching Skills
Identifiers- Ryans Teacher Characteristics Schedule

An experiment was conducted to investigate three orthogonal dimensions of school districts--per pupil wealth, urbanization, and socioeconomic status--and three aspects of schools--socioeconomic status, weighting of teacher behaviors by principals, and type of supervisory organization--in relation to the level or status of and the valuing of two orthogonal sets of teacher characteristics. The first set was comprised of task performance variables, the second of personal-social variables. First, instruments were developed to measure teacher characteristics. Then a sample of approximately 650 teachers drawn from 52 schools in 20 Indiana school systems was studied. Statistical analysis of the results indicated that school districts above the state average in socioeconomic status hold a substantial advantage in attracting and maintaining intermediate teachers superior in task performance, but there is no general relationship between the task performance of primary teachers or the level of personal-social characteristics of primary or intermediate teachers and any major dimension of school districts. Moreover, the valuing of teacher characteristics is almost wholly coordinate to the presence of a supervisory organization in the school district. (A 26-item bibliography, a description of instrument development, and selected instruments are appended.) (Author/SG)

EDO 24634

II
Final Report
Project No. 2579
Contract No. OE-5-10-059

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

DIFFERENTIAL ASSOCIATION OF ELEMENTARY SCHOOL TEACHER CHARACTERISTICS WITH SCHOOL SYSTEM TYPES

Richard L. Turner

Indiana University

Bloomington, Indiana

September, 1968

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION AND WELFARE

Office of Education
Bureau of Research

SP001896

TABLE OF CONTENTS

CHAPTER PAGE

1. Introduction.................................... 1

2. Measures of Teacher Characteristics 5

    Instructional Tasks.................................. 5

    Characteristics Schedule 10

    Revisions of Form E66 14

    Final Task and Scale Reliabilities 25

3. The Selection of School System Characteristics 27

    Systemic Characteristics and the Behavior Modification-selection Model 30

    Systemic Characteristics and the Differential Attraction Model 37

4. Sample Survey Procedures and the Generation of School System Types 40

    The Generation of School System Types 44

5. Procedural Details and Results 50

    Relationships Among the Tasks and Scales 50

    Relationships of Task and Scale Scores to Major Community Social and Economic Variables, with each School District as one Observation 52

    Relationship of Estimated Socio-economic Status of Schools to Teacher Task and Scale Scores, with Teachers as Observations 55

    Relationships of Median Family Income of School Districts and Length of Teacher Service in Districts to Teacher Task and Scale Scores, with Teacher as Observations 57

CHAPTER PAGE

Relationship of Systemic Norms to Task and Scale Scores 64

Relationship of Teacher Task and Scale Scores to Type of Entry into Their System 68

Relationship between System Norms and System Types 79

Relationships among Wealth, Median Education Level and Ratings, with School Districts as Observations 90

6. Discussion and Conclusions 95

Intermediate Teachers, Community Characteristics and Task Performance 95

Primary Teachers, Community Characteristics and Task Performance 98

Primary Teachers, Community Characteristics and Personal-social Characteristics 99

Intermediate Teachers, Community Characteristics and Personal-social Characteristics 103

Model Revisions and Conclusions 104

Applications 108

References 110

Appendix 1. Instrument Development 111

Appendix 2. Selected Instruments 133

LIST OF TABLES

TABLE PAGE

2-1 Relationships between Problems Encountered in Student Teaching and Scales from Form E66 and of the Characteristic Schedule 13

2-2 Intercorrelations Among the Scales in Form E66 of the Characteristics Schedule, N=214 15

2-3 Correlations Among Split-halves of the Scales in Form E66 of the Characteristics Schedule, N=214 15

2-4 Factor Loadings on Three Factors Extracted from the Split-halves of the Scales in Form E66, N=107 17

2-5 Item Overlap between the Scales of Form E66 18

2-6 Split-half Reliabilities of the Newly Derived Scales 19

2-7 Correlations Among the Scales in Form E67 and Factors of the 16 PF 23

2-8 Stability and Split-half Reliability Estimates of the Scales in Form E67 25

4-1 Number of Schools and Teachers Sampled in Relation to Selected Community Characteristics 43

4-2 School and Community Social and Economic Variables 46

4-3 Factors Extracted from 11 School and Community Social and Economic Variables (principal components, varimax rotation to orthogonal factors) 47

5-1 Correlations Among Task and Scale Scores for Primary (above the diagonal, N=282) and for Intermediate Teachers (below the diagonal, N=214) 51

5-2 Factors and Factor Loadings from Task and Scale Correlations for Primary and for Intermediate Teachers 51

5-3 ANOVA for Task and Scale Scores of Primary and Intermediate Teachers, Per Pupil Wealth by Median Family Income, with School System Means as Observations 54

TABLE PAGE

5-4 Distribution of Schools by Socio-economic Status for Wealth by MFI Cells 56

5-4a Distribution of Schools by Socio-economic Status for Median Education Level Cells 56

5-5 Differences in Task and Scores for Intermediate Teachers by Median Family Income in the School District 58

5-6 Differences in Task and Scale Scores for Primary Teachers by Median Family Income in the School District 59

5-7 Differences in Task and Scale Scores for Intermediate Teachers by Years of Experience in their System and Median Family Income in the School District 61

5-8 Differences in Task and Scale Scores for Primary Teachers by Years of Experience in their System and Median Family Income in the School District 62

5-8a Number of Teachers in each Group in Tables 5-7 and 5-8 63

5-9 Mean Scores of Principals on Four TBWS Scales by the Socio-economic Status of their Schools 66

5-10 Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6 + Years in their System and Relationships of Scores to Ratings (norms) in the Hi MFI-Hi Wealth (Type 1) Districts 69-70

5-11 Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6 + Years in their System and Relationships of Scores to Rating (norms) in Hi MFI-Lo Wealth (Type 2) Districts 71-72

5-12 Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6 + Years in their System and Relationships of Scores to Rating (norms) in Lo MFI-Hi Wealth (Type 3) Districts 73-74

TABLE PAGE

5-13 Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 and 6 + Years in their System and Relationships of Scores to Ratings (norms) in Lo MFI-Lo Wealth (Type 4) Districts 75-76

5-14 Mean Scores of Intermediate Teachers New to their Current System, who Transferred to their System and who have always been in their System, by MFI and by District Type 77

5-15 Mean Scores of Primary Teachers who are New to their System, who Transferred to their System and who have always been in their System, by MFI and by District 78

5-15a Number of Subjects in each Cell in Table 5-15 78

5-16 Relationships between TBWS Scores as Norms and Ratings as Norms, Intermediate and Primary Teachers, Hi MFI Districts 81

5-17 Relationships between TBWS Scores as Norms and Ratings as Norms, Intermediate and Primary Teachers, Lo MFI Districts 82

5-18 Mean Task and Scale Scores for Intermediate and Primary Teachers Rated Hi vs Lo, According to Per Pupil Wealth of School District 83

5-19 Mean Task and Scale Scores for Primary and Intermediate Teachers Pooled, Hi vs Lo Ratings by the Socio-economic Status of Schools and by Wealth of School Districts 84

5-20 Distribution of Supervisory Organizations over School Districts, Wealth by MFI 86

5-21 Mean Task and Scale Scores for Teachers Rated Hi vs Lo, Primary and Intermediate Teachers Pooled, in Hi and Lo MFI Districts with and without Supervision 88

5-21a Mean Task and Scale Scores for Primary and for Intermediate Teachers 0-5 and 6 + Years in their District 91

5-22 Effects for Z Total Among Intermediate Teachers, and for W-S and Involvement for Primary Teachers, Wealth x Education Level x Rating 93

LIST OF FIGURES

FIGURE PAGE

6-1 A Modified Teacher Attraction-Behavior Modification Model 109

ACKNOWLEDGEMENTS

The principal investigator is deeply indepted to the students and staff who served on the project in various capacities: To Judy Carefoot who sheparded the Characteristics Schedule through its early revisions; to Velda Stonecypher who organized the systemic data so that sampling could be done; to Patricia Roelke who helped construct the problem-tasks and the schemes by which they are scored, and who helped collect and analyze the data; to Gordon Greenwood who spent many days travelling with the investigator to talk with school system personnel and to test teachers, to Tom Good and Lester Oppenlander who also helped collect the data, and to Wilma Jones who not only served as secretary to the project, but was chief of follow-up operations in the sample survey portions of the project.

The author is also indebted to Professors Ronald Welch, Milton Marten and James Walden of Indiana University and Professor William Powell of the University of Illinois whose suggestions for revisions in the problem-tasks and whose willingness to involve students in performing them were critical to the project. Special gratitude is also due Professor Herbert Kiesling of Indiana University who kept before the author the possibility of economic rather than psychological interpretations of the data, and to Professor Laurence Brown whose suggestions were frequently of great heuristic value.

Finally, the writer wisher to express his deep appreciation to the superintendents, supervisors and principals who generously cooperated in involving their school systems in the project, and to the many teachers who so unselfishly contributed their personal time to provide the data on which the project rests.

X.

SUMMARY

The project investigated three orthogonal dimensions of school districts: per pupil wealth, urbanization, and socio-economic status; and three aspects of schools: socio-economic status, the weighting of teacher behaviors by principals, and type of supervisory organizations, in relation to the level or status of, and in relation to the valuing of, two orthogonal sets of teacher characteristics. The first set was comprised of task performance variables, encompassing skill in teaching reading, arithmetic and science, separated for primary and intermediate teachers, while the second set was comprised of personal-social variables encompassing warmth-spontaneity, classroom organization, educational viewpoint, emotional stability and involvement in teaching as a career.

The sample was comprised of approximately 650 teachers drawn from 52 schools in 20 Indiana school systems.

The results indicated that school districts above the state average in socio-economic status (as indexed by either median family income or median education level) hold a very substantial advantage in attracting and maintaining intermediate teachers superior in task performance, but there is no general relationship between the task performance of primary teachers or the level of personal-social characteristics of primary or intermediate teachers and any major dimension of school districts. The valuing of teacher characteristics is almost wholly coordinate to the presence of a supervisory organization in the school district. Districts with supervisory organizations show differential valuation of teacher characteristics, with personal-social attributes being valued as those associated with the success of primary teachers, with task performance not valued, while among intermediate teachers, task performance is valued, but personal-social characteristics are not valued as characteristics associated with teacher success.

The socio-economic status of individual schools, the way in which principals weight particular teacher behaviors, and the degree of urbanization of the school district were found to be generally unrelated to either the level of teacher characteristics, or the valuing of teacher characteristics. The per pupil wealth of school districts was found to be important primarily as a determinant of which school districts had supervisory organizations, thus, wealth has an indirect rather than a direct influence on teacher characteristics.

CHAPTER I

INTRODUCTION

As a form of work, teaching occupies an ambiguous position between the skilled occupations and the traditional professions. Like the traditional professions, teaching ideally involves substantial latitude for the practioner in interpreting his specific work tasks and in developing individual style and approach in performing them. It is because teaching has this professional side that the characteristics of teachers are important--these characteristics enter into and are presumable important determinants of the way in which the work is interpreted and the style in which it is conducted by individual practitioners.

Unlike the traditional professions, however, teaching, and especially public school teaching, is closely bound to a work organization, the school system, which in turn is inextricably bound to the community which supports it. A teacher does not "practice" in the traditional sense, rather, he is employed, as one is employed in the skilled occupations, and his employer does in fact have a role in interpreting the work, in regulating the style and manner in which it is conducted, and in setting the criteria for work success and evaluating the extent to which the practitioner or employee is successful.

This ambiguity in the occupational position of teaching suggests that if one is to understand the work, he must view it along two quite different dimensions: that of individual variability in the characteristics of the practitioner, and, that of the work organization in which the practitioner is employed.

Teacher Characteristics. Although it is currently fashionable to study "teacher behavior" rather than "teacher characteristics," the difference between the two is procedural rather than substantive. Studies in teacher behavior focus on the actions of the teacher in the classroom and on the interpretation of these actions. Insofar as the actions are consistent from day to day however, observations of them leads ultimately to an induction about the behaviors of the teacher and to statements which interpret these behaviors as characterizing him. For example, under the impetus of Flander's system (7) for observing classroom behavior, teachers become characterized as "direct" or "indirect".

The more traditional approach to teacher characteristics differs from the "teacher behavior" approach in that the behavioral characteristics utilized are either those which have been induced from studies of the behavior of the general population, or some segment of it, or else characteristics formulated in some theory, operationalized, then applied to the general population. After some validity has been established for measures of characteristics developed in either of these two ways, an attempt may be made to examine their relationship to teaching with the hope of showing that certain characteristics identifiable in the general population hold a particular relevance to teaching. An example of the application of characteristics induced from a segment of the general population then applied to teachers may be found in Soar's (18) recent use of characteristics from the MMPI as presage variables in his study of teacher classroom processes. An example of the application of a theoretically based characteristics to teaching lies in a recent study by York (26) in which relations were examined between the scales in Edward's Personal Preference Inventory, based on H.A. Murray's need theory, and beginning teacher problems, largely the writer might point out, withoutsignificant results.

Whether one uses the teacher behavior approach to teacher characteristics or the more traditional ("presage variable") approach to them depends primarily on the kind of risk he wishes to take. The teacher behavior approach minimizes the risk of finding that the characteristics he utilizes are "occupationally irrelevant" to teaching, since he induces the characteristics from observing the behavior exhibited in the performance of the work. This procedure, on the other hand, introduces a kind of theoretical risk in that the characteristics described and utilized may be dissociated from current theoretical viewpoints in psychology. The risk in the traditional approach are exactly the reverse, namely, one uses characteristics of theoretical relevance, thus minimizing this risk, but the risk of finding that these characteristics have no special bearing on teaching is high.

With respect to the two types of risks noted above, the present study sought primarily to minimize the risk that the characteristics examined would be occupationally irrelevant. At the same time, the risk of dissociation with a particular psychological theory was taken.

In developing the measures of the characteristics examined in this study, the writer continued to assume, as in previous studies (20, 21) that teaching, and especially elementary school teaching, involves the interplay of two different sets of characteristics. One of these sets involves the problem-solving skills of the teacher in diagnosing learning difficulties, and organizing or sequencing learning materials. This set of skills is linked to the professional side of teaching in the sense that one of the distinguishing characteristics of a profession is that the practitioners are confronted with highly specific problems for which there are no routine, pre-learned solutions. To use the language of an earlier page, the practitioner must interpret the nature of the task before him and move toward some resolution of it.

The second set of characteristicswas drawn, broadly speaking, from the personal-social domain. It is important, however, not to see these characteristics simply as personality traits. Rather, they represent a diverse set of characteristics including preferences for particular approachs to instruction, elements of personality attributes, and what

Ryans' (17) has labeled "teaching style." These characteristics are distinctly a-theoretical from a psychological viewpoint, having been derived, for the most part, from empirical studies of teacher classroom processes or teacher school-linked behavior.

In sum, aside from the assumption that one expects an interplay between the problem-solving skills of the teacher and his personal-social characteristics, the teacher characteristics entering the study were chosen wholly because they held the promise to be occupationally relevant characteristics, not because they are attached to one or another psychological theory. This is not to say that it is impossible to interpret the relevance of the characteristics to the work of teaching from a theoretical viewpoint, but only to say that the characteristics were not derived from a particular theoretical context in psychology.

The Work Organization. As Edward Gross (9) has suggested, the work organization consists of those work roles, customs, routines, beliefs and values which, in combination, form a system by which some goal, a good or service, is produced. While all societies have work organizations, in an industrial society they are clearly highly organized and complex. With respect to teaching, one can, as Gross has suggested, identify two work organizations. The professional association is one, the school, or more exactly for the public school teacher, the school system, is the other. In this study, the school system, rather than the professional association, is viewed as the primary work organization for the elementary school teacher.

Because the work organization is multi-faceted, it may be examined from more than one perspective. Two such perspectives are employed herein: a psychological, or perhaps a social-psychological perspective, and an economic perspective. Viewed in psychological perspective, the school system as a work organization may be seen as the social unit which has control of much of the reinforcement, positive and negative, available to the practicing teacher. For the beginning teacher, the school system of course controls his very employment, while for experienced teachers some control of salary and of social approval or disapproval, both presumably powerful secondary reinforcers, lie in the authority system. Formally speaking, either supervisory personnel or principals, or in some cases both, have direct control over the reinforcers. It is also true, however, that other teachers, especially senior teachers, may exercise substantial control either by holding the sanctions of the group of practitioners, such as social exclusion, criticism or direct hostility, or by having especially good access to the authority structure, hence to the sanctions it holds. In short, the school district has at its disposal the means by which to influence, and to a degree control, the behavior of the teacher.

From the viewpoint of this study, it is important to note than when one has control of important reinforcing agents, he also has control of the occasions upon which they will be delivered. This is to say that he has control of the criteria which are to be met in order to avoid a negative reinforcement or to gain a positive one. By this mechanism, the school district can, within certain economic limits, develop and maintain a set of criteria concerning what kinds of behavior are "successful" in the school district.

Turning now to the work organization in economic perspective, there are two reasons why this perspective must enter the present study. First, there has been substantial recent work in which the school district is treated as directly analogous to the commercial-industrial firm for which factor inputs-product outputs or production functions may then be generated. By utilizing pupil achievement as the product measure, the extent to which achievement, intelligence and socio-economic status conrolled, is a function of school district wealth, per pupil expenditure, teacher quality, salary, and other input variables common to school systems may then be examined. Such studies are important to the present study because some of them, such as that by Coleman (3)*, indicate that teacher characteristics, even crudely indexed, are associated with a significant portion of the variance of pupil achievement when other variables are controlled, while others, such as those by H.T. James (11) and Kiesling (12) suggest community economic variables holding significant associations to pupil performance, thus providing a clue to variables which may be associated also with teacher characteristics in communities of different types.

Second, from an economic viewpoint there is no escaping the fact that teachers offer a marketable service, and that there is competition among the "firms" for this service. All teachers, however, do not offer the same quality of service, nor do all school systems equally compete for these services since some lie at a disadvantage in wealth, cultural level, location or working conditions. Thus one might anticipate that if an occupationally relevant set of teacher characteristics can be identified, and if these characteristics are associated with the quality of teaching, these characteristics will be differentially distributed across school districts.

Objectives. The study had three major objectives:

1. To continue the development of measures of the problem-solving skills and the personal-social characteristics of elementary school teachers, and to substantiate the occupational relevance of these measures.

---

*The Coleman study was not intended as a production-function study of course, but many of the results of the study appear essentially in this form.

2. To determine the way in which these characteristics are distributed across school districts when they are classified according to the social and economic variables undergirding the communities in which they lie, and according to selected aspects of school district organizational structure.

3. To determine the relationships between the social and economic characteristics of school-communities and the teacher characteristics associated with success in these communities, thus, to learn which teacher characteristics are accorded value in school communities differing in their social and economic characteristics.

In addition to these major objectives, the study also had a number of minor objectives, most of them involving the examination of the level (or status) of teacher characteristics and the teacher characteristics identified with success in relation to sub-clusters of variables. Two principal sub-clusters were employed. One was associated with individual schools, including socio-economic status, the personal characteristics of the principal, and the values placed by principals on certain types of teacher behaviors. The second cluster involved the experience characteristics of teachers, including not only total teaching experience, but also experience in the particular school and school system with which she was associated. The objectives involving these sub-clusters of variables were regarded as minor only in the sense that they were not the central determinants of the design of the study. As determinants of results, however, certain of the variables turned out to be of substantial importance.

The major objectives of the study were accomplished in two separate phases. The first phase, extending from 1964-1966, wholly involved the development and substantiation of the measures of teacher characteristics. While a number of different types of empirical studies were conducted during this phase, all were aimed at testing the validity of the instruments. The second phase, extending from 1966-1968, consisted wholly of designing and conducting a state-wide sample survey of elementary teacher characteristics and their differentiation according to school district types.

CHAPTER 2

MEASURES OF TEACHER CHARACTERISTICS

As observed in the preceding chapter, two different sets of teacher characteristics were of interest in the project. One of these sets deals with the problem solving skills of the teacher in application to her "work" or instructional tasks. The other set deals with her school-linked personal-social characteristics.

The measurement of selected problem solving skills was accomplished by two separate forms of the same instrument, one form for intermediate teachers and one form for primary teachers. The problem-tasks in these forms were structurally parallel, but with the exception of one task, the content of the tasks was adjusted to the particular grade range taught, i.e. primary or intermediate.

The measurement of the personal-social characteristics was accomplished by means of selected items from D.G. Ryan's Teacher Characteristics Schedule, reconstituted into new scales, together with two new scales added by the writer.*

The group of instruments used in the study were given the over-arching title "Behavioral Dimensions of Teaching," and the components were then separately titled, "Instructional Tasks--Primary;" "Instructional Tasks--Intermediate" and "Characteristics Schedule."

In the material which follows, the instructional tasks are first described, then their validity discussed. Next the Characteristics Schedule is discussed in conjunction with four pilot studies in which data bearing on its validity, and its final revision, were obtained.

Instructional Tasks.

The five tasks in each form of this instrument appeared in exactly the same order in each: Tasks 1 and 2 were tasks in teaching reading, Task 3 and 4 were tasks in teaching arithmetic, and Task 5 was a task in teaching science. Compared to similar instruments used in earlier studies by the author, the instruments in question emphasized the subject-matter scope of the teacher's problem-solving skills in teaching rather than his depth of skill in a particular subject. Thus the instruments may be considered appropriate for sample-survey work in which an overall assessment of teacher instructional problem solving skills is of interest, but inappropriateto diagnostic work or to experimental designs in which depth is important.

---

*A detailed discription of the development of the problem-solving tasks and the personal-social scales is given in Appendix 1. Tasks and scales developed during the project but not employed in the sample survey portions of it are also discussed in this Appendix.

Task 1 is identical in both the primary and intermediate forms. In this task a list of 36 Dolch words which have been read on sight by a pupil is presented to the teacher. The pupil made ten errors in the list; the errors made are written in. The problem for the teacher is to group the errors and place an appropriate label on each group. Both the groupings and the labels are scored.

Task 2 is very highly similar in both the primary and intermediate forms. In it the response protocol of a pupil who orally read a passage appropriate to her grade level (5th grade in the intermediate task, 2nd grade in the primary task), together with her responses to comprehension questions over the passage is first observed by the teacher, who is instructed to examine the various word errors and answers to the comprehension questions. Subsequently, the teacher is instructed to examine six reading exercises and four passages with the objective of rating each one on a three point scale in accord with their relevance to obtaining additional pertinent information about the reading skills of the pupil. Separate scores are given for the teacher's ability to handle the passages and her ability to handle the exercises.

Task 3 varies between the primary and intermediate forms more than any other of the tasks. In the intermediate form of this task the teacher examines an arithmetic exercise of mixed problems and examples completed by a pupil at 5th grade level. The pupil's response protocol has been fixed in such a way that he systematically misses problems in some form-classes and randomly errs in others. The teacher's problem is to rate on a three point scale each of ten statements about the errors made in accord with the degree to which she would focus on a particular type of error if a follow-up interview were held with the pupil.

In the primary form of this task, the teacher examines addition and subtraction examples completed by a pupil at advanced second or low third grade level. Again, the response protocol has been fixed with systematic and random errors in different form-classes. Rather than rating statements about errors, however, the primary teacher is instructed to examine ten activities and exercises in arithmetic (many of which are given in pictorial form) and rate each one on a three point scale according to its pertinence as follow-up activity for the pupil, given errors he made. It should be noted that this task is structurally identical to Task 2 (reading) as well as similar to Task 3 in the intermediate form.

Task 4 is structurally identical in the primary and intermediate forms. In the intermediate form the teacher is asked to rank seven division problems in the order of their difficulty for intermediate pupils, while in the primary form she is asked to rank eight addition problems in the order of their difficulty for primary pupils.

Task 5 in both the primary and intermediate forms asks the teacher to examine 25 science-related questions posed by children and then place the questions into groups which might be used as the basis of science units. For the primary teachers, the questions listed were drawn from those asked by primary pupils, while intermediate teachers observed a list drawn from questions posed by intermediate pupils. The basis for scoring this task, which may be of some interest to the reader, is discussed in detail in Appendix 1.

The emphasis on the structure of the tasks in the descriptions above is pertinent primarily because the structure of the task is one determinant of the kinds of intellectual operations required by it. Thus, parallel structure in the tasks for primary and intermediate teachers assures that each group performs the same operations in their tasks. Repetion of structure among tasks in the same form, on the other hand, provides the possibility that a factor attributable to an intellectual operation, such as induction, may be extracted from the tasks independent of their subject matter content. For example, Task 1 and Task 5 both require inductively formed groups, but in divergent subject matters. Task 2 and 3 also require inductions, but the number of complex discriminations to be made both before and after the inductions are made is substantially greater than in Task 1 and 5. If the discrimination-induction operation required in Task 2 and 4 transcends their subject-matter, however, they should show common variance. Task 4 is less clearly an inductive task than the other four tasks, and involves a programming or sequencing skill based partly on discrimination and partly on previously attained concepts.*

Instructional Task Validity. The type of validity appropriate to the tasks in question is construct validity insofar there is no single criterion against which their validity may be ultimately judged. As well as involving the use of multiple criteria, construct validity also embraces the other major types of validity.

The content validity of the tasks presents a complex question. The nub of this question lies in deciding what the relevant universes of content are, and how it is one knows when he has properly sampled them. Three universes are relevant, namely, the universe of content actually taught in elementary schools, the universe of content held by some person or group as that which ought to be taught, and then a quite different universe, that of the operations performed by teachers when they teach the content. The first two of these universes overlap, and it is from the overlapping area that the writer attempted to draw the content of the tasks. Knowledge that the content is drawn from the overlapping area is only probabilitic, however, with the probability guaged, in the present case, by the extent to which the content utilized recurs across the common textbooks in reading, arithmetic and science in the elementary school. While the use of content common to textbooks enhances the probability that a particular teacher does in fact teach that content, the possibility nonetheless remains that in the deviations among school systems and schools there is some school that does not teach the content sampled by the tasks.

---

*A reading task structurally similar to Task 4 was constructed and used in several preliminary studies. The study by Brown (described at a later point) showed that performance this task was negatively related to pupil gains, however, and it was subsequently dropped from the battery.

The separation of the instructional tasks into primary and intermediate forms of course increases the probability that a teacher teaches the content sampled in the form of the tasks he takes, but this probability is far from perfect. For example, a task which is content valid for third grade teachers in arithmetic will almost certainly not be equally valid for first grade teachers. Unless separate tasks are designed for each grade level this problem cannot be fully solved. The means for dealing with it in this study was to balance the grade level of the content across various tasks so that teachers from various grade levels within each grade range had about equal opportunity.

The extent to which the universe of operations performed by elementary teachers is represented in the tasks is unknown because the universe of operations performed by elementary teachers is in fact unknown. It is of course hypothsized that the operations sampled by the tasks are drawn from this universe. Knowledge that this universe has been sampled is not direct, rather, it is inferred from the ability of the tasks to meet the full range of criteria employed to establish the construct validity, the "occupational relevance" of the task.

The concurrent and predictive validity of the tasks varied from task to task. In an earlier study of beginning teachers (21) Task 1 was concurrently associated ($p<05$) with teaching success, as appraised by supervisory personnel, at the end of the second year of teaching experience, but scores taken before teaching began did not predict success on the same criterion. Performance on Task 1 and Task 2 for intermediate teachers, was shown by Brown (2) to be concurrently associated ($p=10$) with pupil gain in reading achievement during a special remedial program given during the summer in a metropolitan area. The relationships between teacher performance and pupil achievement (post-test scores on the Stanford adjusted by analysis of covariance for pre-test scores) was not perfectly linear, however, with teachers scoring in the high average range (5 to 8 points out of 10 points) obtaining the better gains. Two aspects of this study should be further noted. First, Task 2 for primary teachers was not used in the study, hence its validity against the criterion employed remains unknown. Second, while the pupils in the program were of normal intelligence and were drawn from grades4-6, they nonetheless represent a population of pupils who had not adequately responded to normal reading instruction, and may therefore comprise an inadequate group against which to guage the relevance of teacher task performance to achievement gain among typical pupils.

Among beginning teachers Tasks 3 and 4 (arithmetic) for intermediate teachers were shown both to predict ($p<01$) success, as rated by supervisory personnel, at the end of the second year of teaching experience, and to be concurrently associated ($p<.05$) with success on the same criterion. Among experienced teachers the same tasks were also associated ($p<05$) with teacher success, again on the criterion of supervisory rating. In addition, these two tasks were those most strongly associated with teacher performance in a study (20) which showed that pupils passing between pairs of 4th and 5th grade teachers who were high (above the median) in task performance in teaching arithmetic achieved significantly more, by the middle of grade 5, intelligence and previous achievement controlled, than did pupils passing between pairs of 4th and 5th grade teachers who were low (below the median) in task performance.

Finally, Tasks 1, 3 and 4 were the principal contributors to a pooled Z score derived from a total of eight teaching tasks in reading arithmetic which predicted (along with selected factors from Ryans' TCS) which beginning teachers would be viewed by principals as having discipline problems, and problems with setting appropriate levels of expectation for pupils in the elementary grades.

As the foregoing review suggests, the validity evidence available was much greater at the onset of the sample survey phases of the project for Tasks 1-4 in the intermediate form than for any of the remaining tasks. Indeed, that Tasks 2-4 in the primary form and Task 5 in both the intermediate and primary form might be shown to have reasonable validity was predicated largely on their parallelism with tasks for which substantial validity evidence was available.

Characteristics Schedule.

The reconstruction of Ryans' Teacher Characteristics Schedule was carried out in two phases. In the first phase, described in detail in Appendix 1, 230 of the 350 items, and 3 of the 10 scales in the original instrument were eliminated. At the same time, two scales were under construction which were ultimately placed in the Characteristics Schedule. The second phase began with Form E66 of the schedule, which included 120 items from the original TCS and 30 new items. Three major pilot studies* were conducted with Form E66, while a fourth pilot study involved only one of the new scales placed in Form E66.

---

*These studies were pilot studies only in the eyes of this investigator. For the graduate students who conducted them, they were dissertations.

The Quinn Study (14). This study focused on two questions: 1) whether there were differences in the personal-social characteristics of teachers in inner-city schools and teachers in fringe area middle class schools, and 2) whether these characteristics were differentially valued in the two types of schools as indicated by the characteristics of teachers rated high versus those rated low by principals in these schools. The occasion for these questions stemmed from the results of an earlier study of beginning teachers by the writer which indicated that the scales on the original TCS predicted who would be viewed as successful in systems composed largely of middle class schools, but did not predict who would be viewed as successful in systems composed largely of working class schools, although score levels in the two types of systems did not differ. The problem in the latter study lay in the confounding of the socio-economic status of schools with the type of system in which they were located. This difficulty was corrected in the Quinn study.

The study was conducted in a large metropolitan area, from which seven inner-city elementary schools, having 102 teachers, and nine middle class elementary schools, having 111 teachers, were drawn. The teachers completed Form E66 of the Characteristics Schedule, and principals subsequently ranked the teachers in four quarters (forced rating) according to their over-all success. In addition, several items of information, including age, experience in the system and the like were obtained from teachers.

The Scales appearing in Form E66 were as follows:*

X--Warm, friendly understanding versus cool aloof (28 items)
Y--Responsible, business-like versus shipshod (28 items)
Z--Stimulating imaginative versus dull, routine (24 items)
R-R_1--Attitude toward pupils (34 items)
Q--Attitude toward school staff (30 items)
B--Child-centered versus subject-centered (30 items)
S--Emotional adjustment (30 items)
TI--Teacher Involvement (14 items)
ACS--Authority centering versus authority sharing (16 items)

Among these scales the first seven were derivatives of the original TCS, while the last two were experimental scales, the final one of which (the ACS) was designed to be added to the B scale.

---

*The scale labels follow those established by Ryans, except for TI and ACS, which are new scales.

The results of the study were quite similar to those in the study of beginning teachers. There were no differences in score levels between types of schools, but teachers rated high in middle class schools showed significantly higher scores on scales X (warm, friendly), scale Z (stimulating, imaginative) and TI (Teacher Involvement) than did teachers rated low. In inner-city schools there were no differences on any scale between teachers rated high and those rated low.

Although these results were generally quite satisfactory, two difficulties remained with Form E66; the first was the very great overlap of items between the scales, a problem discussed in detail in Appendix 1. The second was shrinkage in the reliability coefficients for the revised scales compared to those obtained from preceding sub-studies using the same scales with graduate students, and shown in Appendix 1. A massive revision of the structure of the Schedule was therefore once more undertaken in an effort to increase the homogeneity of the scales and decrease overlap between them. In the meantime, however, Yoder initiated his study with Form E66, and Mills her study utilizing the Teacher Involvement scale. These studies will therfore be described before discussion of the revision of E66.

The Yoder Study (25). The question in focus in this study was whether the scales in the Schedule could be used to predict the occurence of teaching problems in student teaching among elementary and secondary preparatory teachers attending a small, church-related college in Indiana. Like the Quinn Study, the questions posed in this study rested in part on the results of an earlier study by the writer (22) in which certain TCS scales were shown to predict the occurence of problems in discipline, teaching reading, and setting pupil expectancy levels among beginning teachers.

The 108 student teachers involved in this study completed the Schedule in late August, then went immediately to their student teaching assigment. In October, Yoder sent to each supervising teacher a problem rating sheet of his own design, covering the following areas: knowledge of subject-matter, lesson planning, pupil motivation, classroom management, teacher-pupil rapport, pupil evaluation, teacher confidence, and interpersonal school-staff relationships. Within each of these areas numerous items explicating the problem area were given, and the teacher checked off the occurence of the problems stated in these specific items as they occured between October and January, returning the problem rating sheet at the end of the semester.

The results of this study were again quite satisfactory. The problem areas and the Schedule scales significantly predicting them are shown in Table 2-1.

TABLE 2-1. Relationships between Problems Encountered in Student Teaching and Scales from Form E66 and of the Characteristics Schedule.

		SCALE			
		X	Y	Z	Q
PROBLEM AREA	Classroom management	F=7.44 (.01)	F=25.47 (.001)	F=16.78 (.001)	F=13.83 (.001)
	Pupil motivation		F=4.18 (.05)		
	Teacher-pupil rapport		F=6.94 (.01)		F=4.78 (.05)

Compared to the earlier study by the writer, the ability of scales Y, Z and Q to predict classroom namagement problems was considerably enhanced in the study by Yoder, while the discrimination ability of the X scale continued at about the same level. Beyond this point, the problem categories in the two studies are not fully comparable. Nonetheless, it seemed apparent that a drastic shortening of X, Y, Z and Q scales had not decreased their validity, and had perhaps increased it, at least on the criterion of the prediction of classroom problems. It should be noted that Yoder did not employ either the Teacher Involvement Scale or the ACS in the analysis of his data, hence, results for these scales were not available.

The Mills Study (13). In this study, the concurrent relationships of the Career Motivation Scale (CMS) and the Teacher Involvement Scale (TIS) to selected aspects of career development patterns were examined among 326 female elementary education graduates, three to five years after graduation. Although the study involved the CMS, a scale not used in the project, as well as the TIS, the results involving both of these scales, and other factors, are important to the interpretation of the validity of the Involvement scale.

Among Mills' findings were that whether a teacher remained in or dropped out of teaching during the first 3-5 years depended on ($p<.001$ in all cases) husband's salary, number of children, and CMS score level. TIS score level, however, was not related to whether the teacher remained in or dropped out of teaching. TIS scores were nonetheless positively related to several other factors, namely, definite plans to return to teaching versus no plans ($p<005$); plans to remain in teaching versus plans to quit ($p<001$);

satisfaction versus dissatisfaction with teaching $(p<.001)$; having a mother who was a teacher versus having one who was not $(p<05)$; having husband in a lower middle class or working class occupation versus a middle class occupation $(p<.05)$; having one or more children versus having no children $(p<05)$; teaching at a single grade level versus shifting grade levels $(p<05)$; teaching special classes or substitute teaching $(p<05)$ and reporting that the preparatory teacher education program was adequate versus reporting it to be inadequate $(p<05)$. An additional finding of importance was that the CMS was uncorrelated with the TIS.

The fact that the CMS discriminated those who remained in teaching versus those who dropped out, while the TIS did not make this discrimination, seemed to damage the validity of the TIS. This result came about in a logical way, however. When all graduates are considered, there is a positive relationship between TIS scores and having one or more children. In the same group, there is a negative relationship between remaining in teaching and having one or more children. Because these two relationships run in opposite directions, it is unlikely that TIS scores will be associated with remaining in teaching since a substantial portion of those with higher scores drop out for reasons of pregnancy. If the group is changed to include only those currently engaged in teaching, the proper relationship does appear, namely, those who plan to quit teaching show much lower Involvement scores than those who plan (presumably until pregnancy catches up with them) to remain in teaching.[*]

The results associated with the TIS in the study by Mills, together with the results reported in Appendix 1 (that instructional task score increases among undergraduates were closely associated with their TIS scores) strongly suggested that Involvement is an extremely important teacher characteristic, and the decision was made to add it as a permanent part of the Characteristics Schedule.

Revision of Form E66.

As noted in the discussion of the Quinn study, Form E66 of the Characteristics Schedule suffered two serious deficiencies. The first was low reliabilities and the second, a serious overlap of items between scales. To grasp the nature of the difficulties with the reliabilities, it is informative to observe Table 2-2 and 2-3. Table 2-2 gives the intercorrelations among the scales, and Table 2-3 correlations between split-halves of the same scale.

---

[*] There is a positive correlation between TIS scores and scores on Scale X (warm, friendly). In an earlier study of beginning teachers a tendency for high scorers on Scale X to become pregnant was also observed.

TABLE 2-2. Intercorrelations Among the Scales in Form E66 of the Characteristics Schedule, N=214.

Scales	Y	Z	R	Q	B	S	TI	ACS
X	33	72	67	49	-43	64	31	-37
Y		31	44	64	-05	43	18	-10
Z			45	43	-26	44	32	26
R				74	-28	68	24	28
Q					-15	64	18	19
B						-29	-15	52
S							18	-37
TI								-09

TABLE 2-3. Correlations Among Split-halves of the Scales in Form E66 of the Characteristics Schedule, N=214.

Scales	Xb	Yb	Zb	Rb	Qb	Bb	Sb
Xa	34						
Ya		20					
Za			25				
Ra				53			
Qa					52		
Ba						41	
Sa							44

Comparison of the correlations between halves of the same scale with correlations between scales presents a truly incredible state of affairs. For example, the correlation between scales X and Z is .72, but the correlation between the halves of X is only .34, and the halves of Z .25. The correlation of Y with Q is .64, but the correlation of the Y halves is only .20, although the correlation between the Q halves is a reasonable .52. This kind of situation arose from the fact that the items in each scale were relatively heterogeneous, which undoubtedly enhances the validity of the scales since the criteria against which validity is appraised are heterogenous, but the same item scored the same way appeared in more than one scale. For example, of 39 items in scales Y and Q pooled, 13 were common to both scales and were scored the same way in each scale. Thus, only 26 of the items between these scales had any opportunity to be correlationally independent.

The initial effort to correct this problem was a second* attempt to factor analyze a matrix of phi coefficients generated from correlating each item with all the others. A coherent factor structure could not be obtained. The assumption was then made that the individual item reliabilities were too low for a factor analysis of individual items, and factor analysis of the split-halves plus the ACS scale was conducted. The results, using a principal components method with varimax rotation to orthogonal factors, were distinctly interpretable. The three factors extracted and the loadings on each are shown in Table 3.

The three factors extracted clearly represented scales Y, Q, R and S in one set; scale B in the second set; and scales X and Z in the third set. The first of these factors appeared to represent a general factor. On the basis of the Yoder study, as well as the previous study of beginning teachers, however, scales Q and Y but not scales R and S were known to predict problems with classroom management. Thus the general factor was split into two subsets on the assumption that the validities of the two sub-sets of scales within it were not equal, even though both sets loaded on the same factor. The general factor was, however, also retained for experimental purposes.

At this point, the second deficiency in Form E66, the overlap of items between scales, remained. The extent of this problem may be viewed in Table 2-4, in which the number of items appearing on any one scale that also appeared on some other scale may be observed.

---

*The first attempt is mentioned in Appendix 1.

TABLE 2-4. Factor Loadings on Three Factors Extracted from the Split-halves of the Scales in Form E66, N=107.**

VARIABLES	FACTORS* 1	2	3
ACS		-80	
Xa	46		66
Xb		44	61
Ya	61		
Yb	63		
Za			70
Zb			76
Ra	72		
Rb	71		
Qa	86		
Qb	71		
Ba		-63	
Bb		-78	
Sa	71		
Sb	49		

*Factor loadings of less than .40 are excluded.

**Only half the subjects were used in the analysis, with the other half retained for use in a cross-validiation study.

TABLE 2-5. Item Overlap Between the Scales of Form E66.

Scale	Total items	Scale						Total overlap*
		Y	Z	R	Q	B	S	
X	28	2	13	10	3	1	7	36
Y	24		2	5	10	2	5	26
Z	24			4	2	1	2	24
R	34				9	0	10	38
Q	28					0	7	31
B	30						0	4
S	30							31

*The total overlap can exceed the number of items in the scale since a single item may appear on several scales.

The solution to the overlap problem was simpler than one might infer from observation of Table 2-5. The conjoining of scales, for example, Q and Y, removed a very great proportion of the overlap, while the remainder was removed by examining the biserial r's for the items on each scale, then ordering the overlapping items to that scale with which they were most closely correlated.

After the scales were re-organized, as described above, they were renamed. The conjoined scales X (warm, friendly) and Z (stimulating, imaginative) were renamed "Warmth-spontaneity;" the conjoined scales Y (organized, business-like) and Q (attitude toward school staff) were renamed "Organization;" the conjoined scales S (emotional stability) and R (attitude toward pupils) were renamed, "Stability;" while the conjoined scales B (educational viewpoint) and ACS (authority centering versus authority sharing) were renamed "Viewpoint." Low scores on the latter scale continued to be associated with a child-centered viewpoint, high scores, a subject-centered viewpoint. Thus, the seven scales from Form E66 were reduced to four scales.

Subsequently, each of the four new scales and the scales in the general factor were split into random halves, and the reliability of each scale run using the scores of that half of the subjects in the Quinn study who had not been used in the factor analysis. The results appear in Table 2-6.

TABLE 2-6. Split-half Reliabilities of the Newly Derived Scales.

	Scale				
	Warmth-spon.	Organiz.	Vpt.	Stab.	General
rxx	.66	.60	.58	.78	.57

While the reliability coefficients appearing in Table 2-6 were by no means as high as the writer would have liked, they represented a substantial improvement over those for the old scales. Moreover, the fact that rxx for the general factor was lower than for its component scales distinctly suggested that separating the Y-Q components and the R-S components in this factor had resulted in better reliabilities than leaving the general factor intact. The viability of this separation was also borne out in a subsequent study.

Following the revision of the scales in Form E66, another study was launched jointly between CRP 2579 (Turner) and CRP 6-8235-2-12-1 (Denny) in which the relationships between teacher characteristics, teacher classroom behavior and pupil changes in creativity were examined.

The Turner-Denny Study (23). This study fell into two principal parts. In the first part, conducted wholly by Denny, observed teacher classroom behaviors were related to changes in creativity scores among sixth grade pupils. In the second part, conducted by both Denny and Turner, the teachers characteristics assessed by the revised scales in Form E66 were related to both the observed classroom behaviors of teachers and to changes in pupil creativity.

Of the five measures of creativity adapted by Denny from those constructed by Guilford and others (10), four were found to be sufficiently reliable for use. These four were Redefinition (unusual uses for common objects), Fluency (having many ideas), Flexibility (other uses for common objects), and Sensitivity (ability to think of problems associated with common situations or objects).

It should be noted that Redefinition differs from Flexibility primarily in that the former is verbal while the latter is pictorial. These measures were administered to 30 sixth grade classes in October and again in April. The resulting dependent variable was the April scores adjusted by analysis of covariance for the October scores and for intelligence.

The observational schedule used by Denny was highly similar in format to the OScAR, but contained items held by Denny to be especially relevant to pupil creativity change as well as items intended to describe the general classroom behaviors of the teacher. There were ten items in the Schedule, as follows:

Motivational climate--use of threatening versus positively reinforcing stimuli;

Pupil interest--pupil eagerness and attention versus reluctance and irritability;

Teacher-pupil relationship--teacher is attentive to pupil remarks, responds positively, etc., versus teacher interference, abruptness, and the like;

Pupil-pupil relationship--pupils refer positively to success of others, take responsibility, accept individual differences versus teasing, reluctance to share responsibility, etc.

Pupil-initiative--teacher dominates classroom activity versus pupils dominate classroom activity;

Teacher approach--teacher paces lesson to build interest, has materials ready, concludes lesson while interest is high versus the opposite;

Adaptation to individual differences--numerical index of the number of individuals to whom the teacher offers individual attention, times the number of times the teacher differentiates, by the number of minutes observed;

Variation in materials and activities--number of different materials and activities used during total observation period;

Encouragement of pupil divergent thinking--teacher encourages divergent versus convergent thinking;

Encouragement of unusual pupil responses--tally of reinforcements given for unusual (divergent) responses.

Except as indicated above for specific items, the scoring of the items is accomplished by means of the observor assigning weights (positive to negative).

Each teacher in the study was observed three times, each time by three observors. The score used for each teacher was the mean of each item over observors over times observed. As elsewhere discussed by Denny (4), the items in the schedule are surprisingly reliable.

The four revised scales of Form E66 were related to both the observed items of teacher behavior and to the pupil creativity change by analysis of variance with each classroom serving as one degree of freedom.

The results indicated that Warmth-spontaneity was associated with pupil increases in Redefinition ($p<01$) with Motivational climate ($p<.10$), Pupil-pupil relationships ($p<.10$) and with Encouragement of Unusual Responses ($p<.05$), which were congruent with Denny's finding that Pupil-pupil relationships and Motivational Climate were significantly related to increases in Redefinition. Over-all, these results suggested that Warmth-spontaneity is associated with positive reinforcing teacher classroom behaviors, certainly including the reinforcement of divergent pupil responses.

Scores on the Organization scale, on the other hand, were negatively related to pupil increases in Fluency ($p<.02$) but positively related to Motivational Climate ($p<10$), Pupil Interest ($p<05$) and Teacher-pupil Relationship ($p<.01$). This result was congruent with Denny's finding that Teacher-pupil relationship was negatively related to Fluency increases among pupils. These results, together with the knowledge that the Y and Q, scales from which the Organization scale was derived predict an absence of discipline problems in the classroom, strongly suggest that higher scores on the Organization scale are rather closely linked to firm classroom management-control procedures. Such a classroom appears to look good to the observer, but nonetheless leads to a decrease in pupil Fluency. It is quite thinkable, however, that a classroom of this kind leads to quite acceptable increases in convergent achievement.

The results for the Viewpoint scale indicated that the more child-centered teacher (lower Viewpoint scores) maintained a better teacher approach ($p<10$), did more adaptation to individual differences ($p<05$), used a greater variety of materials and activities ($p<05$) and had better general structuring for learning (a combination of items in the Denny schedule) ($p<01$), and obtained greater increases in pupil Flexibility ($p<05$) than did the more subject-centered teacher. Denny found Teacher Approach, but not the other items, to be associated with Flexibility increases. In the main, the results for the Viewpoint scale suggest that this scale is in some degree a measure of teacher flexibility, as indexed by adaptation to individuals and use of a wider variety of materials and activities.

Finally, the Teacher Involvement scale showed a quite strong relationship ($p<.005$) with the encouragement of pupil initiative, but this relationship was inverse, namely, involvement leads to greater teacher domination of classroom activity. Perhaps this is what one would expect from a person highly involved in her work. Nonetheless, there was a positive relationship between Involvement and pupil increases in Redefinition. ($p=.07$)

Over-all, the results of the Turner-Denny study was interpreted to be quite satisfactory, and the final revision of the Characteristics Schedule was then made.

Characteristics Schedule, Form E67. Form E67 of the Schedule represented only a slight revision from Form E66. Six items, all of which were found to be malfunctioning when the new scales were devised, were dropped, while 16 items were added. All 16 items were from a verbal induction scale, the development of which is discussed in Appendix 1. The movement of the Induction scale to the Schedule was done primarily to decrease the face to face testing time with teachers which was approximately one-half hour beyond the one hour and fifteen minute testing time alloted to it for the sample survey. Subsequently, all items were clipped and placed in a box, thoroughly rotated, then drawn out one by one and placed in the Schedule, producing a new, random sequence totallying 160 items.

The Pugh Study. This study, by Richard Pugh, Associate Professor of Education at Indiana University, was conducted while the sample survey was in progress. It is reviewed in this place, however, because it represents an important validity-reliability study of Form E67 of the Schedule conducted quite independently of the principal investigator. Only the validity phases of this study are discussed in this section, while the reliability data are discussed in the next section.

The validity data in the study were comprised of the intercorrelations of the six scales appearing in the Characteristics Schedule and the 16 factors from Cattell's 16 PF among 75 unselected experienced elementary teachers taking the M.S. in Education. The importance of these data lie centrally in the extent to which the CS scales contain elements of personality characteristics commonly found in the general population, as opposed to elements of attributes which might be considered primarily characteristic of teachers. The correlations are shown in Table 2-7.

A distinct constraint in interpreting the relationships shown in Table 2-7 lies in fact that the factors in the 16 PF are themselves empirically derived, are perhaps not themselves fully validated, and have an oblique factor structure. Certainly there is a question in the

TABLE 2-7. Correlations among the Scaled in Form E67 and Factors of the 16 PF

Personality Factor#	W-S	Org.	Vpt.	Stab.	TI	Ind.
Reserved, detached vs outgoing, warmhearted	.45**	.15	-.13	.25*	.13	-.16
Less intelligent vs more intelligent	.17	.10	-.25*	.28*	.21	.40**
Affected by feelings vs emotionally stable	.33**	.25*	-.23*	.33**	.17	.46**
Humble, mild vs assertive, independent	-.15	-.12	-.02	.14	-.19	.09
Sober, serious vs happy-go-lucky, gay	.29**	-.01	-.13	.17	-.05	-.12
Expedient vs conscientious, rule bound	-.03	.15	.12	.11	.15	-.13
Shy, restrained vs venturesome, spontaneous	.30**	-.01	-.29**	.35**	-.01	.06
Tough-minded vs tender minded	.07	-.21	.00	-.21	.08	-.14
Trusting vs suspicious	-.18	-.26*	.11	-.16	-.08	-.14
Practical vs imaginative	-.02	-.18	.01	.01	.03	.03
Forthright vs shrewed	.00	.18	.07	.12	.17	.07
Placid, confident vs worrying, troubled	-.22**	-.32**	.15	-.43**	-.08	-.23*
Conservative vs experimenting	.18	-.03	-.21	.10	.07	.12
Group-dependent vs self-sufficient	-.30**	-.11	.11	-.20	.05	.06
Undisciplined vs controlled	-.19	-.08	.09	-.11	-.05	.02
Relaxed, tranquil vs tense, overwrought	-.16	-.05	.10	-.16	.00	-.23*

#The low score discription is first.
*$p < .05$
**$p < .01$

mind of the writer whether a factor analysis of the correlations would in any sense increase the interpretability of the matrix or enhance in any way the construct validity of the scales in the Characteristics Schedule. Nonetheless, the results are suggestive and provide interesting supplementary data bearing on the over-all validity of the scales.

One may note first that there is a strong element of emotional stability in both the Warmth-spontaneity and the Stability scales. To the extent that these scales are differentiated (the r between them in the group sampled was .55), Stability is perhaps more closely linked to emotional adjustment factors such as freedom from anxiety (i.e. placid, confident versus worrying, troubled) while Warmth-spontaneity is more closely linked to outgoing, warmhearted and group-dependent behaviors. The Organization scale follows similar lines, with such differentation as occurs apparently linked to a somewhat greater number of elements of trusting behavior in this scale than in Warmth-spontaneity and Stability. Again, Viewpoint, and especially child-centerdness, shares elements with many of the same scales as do Warmth-spontaneity and Stability. In both the Viewpoint and the Organization scales, however, the associations with the _16 PF_ are fewer than in W-S and Stability, implying greater unique variance.[*]

Of the remaining two scales, the Involvement scale shows complete independence of the factors in the _16 PF_, and may be interpreted as a measure independent of common personality characteristics. The Induction scale, on the other hand, holds a strong, but not unprecedented, set of relationships to the factors of the _16 PF_. Among the relationships, that between Induction and Cattell's intelligence scale is congruent with the expected relationship between intelligence and induction. The correlation between induction, emotional stability, placidity and tranquility however, indicate a distinct relationship between personality factors and induction. Ironically, the Verbal Intelligence scale originally included in Ryans' TCS was deleted by the writer in the initial stages of the revision of the TCS because it correlated rather closely with the personal-social scales. These correlations, together with the fact that Ryans' Verbal Intelligence scale did not correlate with instructional task performance, while such performance did correlate with the Ohio State Psychological Examination, was taken as evidence of invalidity. It not seems clear, however, that attempts to create an induction scale independent of personality characteristics were quite unsuccessful.

---

[*] Professor Pugh will discuss the matter of the variance of each CS scale independent of the 16 PF scales in a paper planned for _Educational and Psychological Measurement_.

While the basis of the associations between induction and certain personality factors is not clear, the fact that induction is correlated with child-centeredness (r=.33, p<.01) in the same sample of subjects, together with the fact that child-centeredness undergoes drastic changes with student teaching (see the next section) suggests the possibility that the acquisition and/or retention of certain types of personality traits, traits perhaps more common to teachers than the general population, does in fact depend on certain aspects of intelligence such as induction or concept attainment.

Over-all, the relationships between the scales on the Characteristics Schedule and those in the 16 PF suggest that W-S and Stability are most closely linked to personality characteristics common to the general population, while Organization and Viewpoint are considerably less related and Involvement unrelated.

Final Task and Scale Reliabilities.

A second aspect of the study conducted by Pugh involved examination of reliability estimates for the scales in Form E67, including both stability and equivalence estimates. These estimates are considered the final ones in the study partly because they were obtained independently of the principal investigator, and partly because they were obtained under conditions which tested the robustness of the reliabilities.

The stability of the scales was examined among preparatory elementary teachers by testing them just prior to student teaching and again after student teaching, a period of about five months. Prior to the scoring of the second of the two testings, the biserial correlations between each item and the total score for the scale in which it appeared, and the percent passing each item were calculated. Two halves of each scale were then produced by matching the items on bis r and percent passing, subsequently, the two matched halves were scored for the sample of unselected graduate students who had participated in the study involving the 16 PF. The results may be observed in Table 2-8.

TABLE 2-8. Stability and Split-half Reliability Estimates of the Scales in Form E67.

	Scales					
	W-S	Org.	Vpt.	Stab.	TI	Ind.
Stability (N=73)	.56	.43	.45	.63	.68	.29
Split-half (N=75)	.65	.57	.70	.72	.74	.44

Comparison of the split-half coefficients to those appearing in Table 2-6 indicates that using independent samples and a slightly revised form of the Schedule nonetheless yielded quite comparable reliability estimates. The shrinkage appearing in the stability coefficients for Organization and Viewpoint was accompanied by changes in the mean scores on these scales in conjunction with the student teaching. In the case of the Organization scale, the mean increase was approximately one-third of a standard deviation, resulting in a F ratio of 7.70 significant beyond p=01. For the Viewpoint scale the mean increase was one and a quarter standard deviations, resulting in an F ratio of 59.42, significant much beyond the .001 level. In each instance, a score increase may be interpreted to mean a distinct movement toward greater conservatism with student teaching experience.

Both the stability and split-half estimates for the Induction scale were symptomatic of the difficulties encountered in constructing a short scale to measure this attribute, and predictive of its general utility in the sample survey.

The computation of split-half reliabilities for the instructional tasks was completed at the conclusion of the sample survey using the final form of the instruments. The reliabilities of the tasks for intermediate teachers were: reading .73, arithmetic .71, and science .77; and for primary teachers: reading .70, arithmetic .68, and science .83.

CHAPTER 3

THE SELECTION OF SCHOOL SYSTEM CHARACTERISTICS

Just as it is true that some, but not all, of the characteristics of teachers may be regarded as occupationally relevant, so is it true that some, but not all. of the characteristics of school systems may be found to account for differences between systems in teacher characteristics. To locate these systemic characteristics some type of strategy is needed. In the project, this strategy was comprised of two alternate sets of hypotheses and one set of predominately empirical considerations. As suggested in Chapter 1, the two sets of hypotheses, converted to loosely structured models, were on the one hand identified with social-psychological considerations, and on the other, "economic" considerations. The predominantly empirical matter was associated with the actual intercorrelations among certain of the variables chosen.

The social-psychological considerations are identified with a behavior modification-teacher selection model. There are five broad components in this model; 1) community social and economic variables, 2) systemic norms for teacher characteristics (T.C.), 3) a delivery mechanism, 4) the differential reinforcement of teacher behavior, or the differential selection of teachers, and 5) differences between systems in teacher characteristics. In diagrammatic form, this model is as follows:

Community Social and Economic Variables	- -	Systemic Norms for T.C.	- - - -	Delivery Mechanism	- -	Differential Reinforcement of Behavior	- - - -	Differences in T.C. between Systems
					- - - -	Differential Selection of Teachers	- - - - - -	

The "economic"* considerations are identified with a differential attraction model. This model has four major components: 1) community social and economic variables, 2) economically, or socio-economically controlled systemic policies, 3) differential community attractiveness, and 4) the prediction of teacher behavior characteristics-systemic norm congruence. This model may be diagrammed as:

---

*This model is not intended to be an economic model in any formal way; the label "economic" is attached primarily to suggest that the major variables involved are usually associated with economics rather than education or psychology.

Community Social and - - Economic Variables	Economically Controlled - - - Policies	Predicted Behavior norm - - - - - - - - - - - - Congruence	To Differences - - - - - - - Between Systems
- - - - - -	Differential Community Attractiveness - - - - -		

The two models above are in no sense mutually exclusive. Both may operate in any school system, and presumably do operate. In a longitudinal study of the influence of school systems on teacher characteristics, both models might be combined into a single model, since the differential attraction model is essentially the antecedent to the modification-selection model. Nonetheless, in the project they were kept separate, partly because they represented alternative ways of looking at and interpreting the data, and partly because the social-psychological model was generated before the project began, while the economic model was fully perceived as an alternative only after the project was substantially underway.

The components of the modification-selection model were derived predominantly from the results of the preceding project (CRP 1262). Most of these results are given in Appendix 1, but certain of them, as reviewed below, are important to understanding why the modification-selection model is as it is.

The study was conducted among 13 Indiana school systems in which approximately 200 beginning primary and intermediate teachers were pretested with Ryans' TCS and pre and post tested with problem-solving tasks in arithmetic, and in reading, with approximately a two year interval intervening. In addition, numerous systemic variables were observed, including selected community social and economic variables, amount of supervision received by each beginner, and his success as rated by supervisory personnel. The results were:

1. Task performance and personal-social (TCS) characteristics were evenly distributed across districts when teaching began.

2. Increases in problem-task performance among intermediate beginning teachers occurred only in districts (Type A districts) with above average per pupil wealth, a high ratio of working class to middle class pupils, and the presence of a supervisory staff.

3. Within the districts cited above, the degree of change was a linear function of the amount of supervision received.

4. In districts (Type B districts) with below average per pupil wealth, a high ratio of middle to working class pupils, and typically the absence of a supervisory staff, there was no significant increase in performance, and amount of supervision received was irrelevant.

5. In Type A districts, beginning teacher success was predicted by problem-task performance, but not by TCS variables, while in Type B districts, TCS variables predicted success, but problem-task performance did not.

6. In Type B districts, the amount of supervision received was a function of problem-task performance and TCS scales B (child-centered vs subject-centered and Z (stimulating-imaginative vs dull, routine), but no tasks or scales were associated with the amount of supervision received in Type B systems.

These results were interpreted to mean 1) that teacher task performance can be (and is) modified by the school system, 2) that supervision is one school mechanism by which performance is changed, 3) that differences in performance change and differences in the characteristics with which success is associated indicate differences in the norms (or "value structure" or "criteria") for teacher behaviors of characteristics in school systems of different types, 4) that there is some evidence that the norms have a hierarchiacal order, 5) that the norms were associated with, if not a true function of, economic and social variables in the community, and 6) that if the differential modification of performance were continued for long periods of time, experienced teachers in different types of systems would have quite different characteristics.

It is apparent that the main components of the modification-selection model came from these conclusions. There are certain components of the model which did not arise from these conclusions, however. First, differential reinforcement of teacher behavior is assumed to be the particular mechanism underlying changes in teacher behavior, but there was no direct evidence for this point in the study. Second, there was no clear-cut evidence in the study that differential selection occurred, although there was very slight evidence ($p=.05$) that the more child-centered teachers left Type A systems while the more subject-centered left Type B systems. This evidence maintained the possibility that over long periods of time differences in teacher characteristic could come about by selection.

To this same point a word should be said about why the differential attraction model was not viewed as viable early in the project, namely, there was no significant evidence that any one system attracted beginning teachers different in their characteristics from those attracted by any other system. Indeed, the alternative model was not fully seen until a second follow-up of the beginning teacher sample was made at the end of four years, at which time it was noted that approximately 70 percent had left the system in which they were orginally employed. While the latter figure includes both mobility within the profession and attrition from it, it nonetheless made clear that the ability of a system to attract mobile experienced teachers could be a highly significant factor contributing to differences in teacher characteristics between systems, especially if there were differential attraction and/or screening capabilities in school systems of different types.

Systemic Characteristics and the Behavior-Modification-selection Model. Among the components of this model that which holds a key position, but is most difficult to operationally define, is "systemic norms for teacher characteristics." Such norms may best be viewed as a hypothetical construct, hence not directly open to observation, but nonetheless meaningful as inferentially entities which may be assessed through indicators or signs. In the study of beginning teachers, the indicators used to infer the presence of systemic norms were 1) the outcomes of the ratings of success, and 2) changes in teacher performance. Of these two indicators, the outcomes of ratings are viewed as particularly strong, since a rating of teachers forces the system, or persons in the authority structure of the system, to select those persons most valued, as opposed to least valued, as successful in the system. Examination of the characteristics of these opposing groups of teachers then suggests which teacher characteristics are given weight in the criterion or norm employed by the rater. It should be noted that in order to infer the criterial or normative characteristics from this procedure it is critical for the investigator to leave an open field for the rater, i.e. the investigator cannot specify which characteristics he considers criterial, either explicitly, or implicitly in the rating form. Rather, the criterion defining "success" is left unspecified and is inferred from the choices of the rater.

The use of changes in performance as an indicator is supplementary, but important. When the behavior of beginning teachers in a particular group of systems changes in the direction of the norm inferred from the ratings taken in these systems, evidence to support an inference that there is a unitary source, i.e. a norm, is added.

The difficulty with using indicators of the sort described above is not that they are weak, but that they are expensive and cumbersome procedures by which to define the existence of norms. An inexpensive, rapid means of directly testing for the presence of norms for selected teacher characteristics was therefore sought in the project, resulting in a paper and pencil instrument called the Teacher Behavior Weighting Scale.

The Teacher Behavior Weighting Scale (TBWS) was developed not only on the assumption that certain types of norms characterize school systems, but that these norms have a non-ipsative structure, i.e. that there is a universe of norms for teacher behavior present for each person in each school system, and that the same elements are present in this universe, but that the elements differ between individuals in the weights assigned to them. Thus, greater value is placed on some teacher behaviors than others, with the more valued or more heavily weighted members being the dominant members in the heirarchy, and the less valued, less heavily weighted members being non-dominant, but still in the hierarchy. Under this comceptualization, a systemic norm means that there is some agreement within a school system about which teacher behaviors have the heaviest weights, i.e. the ordering of the elements within the universe is relatively homogeneous between persons, especially persons in the authority structure of the system.

The hypothesis that there is an ordered structure in the norms for teacher behaviors or characteristics was taken by inference from a pattern of results appearing in Type A districts in the study of beginning teachers. This pattern was: 1) the degree of task performance increase was a function of the amount of supervision received, but 2) the amount of supervision received was primarily a function of the personal-social characteristics (TCS scales B and Z) of the teachers, while 3) task performance was the central criterion in teacher success. Given the knowledge that the task performance and personal-social characteristics of beginning teachers are orthogonal, this pattern of results suggests that task performance was the criterial or dominant member of the hiearchy in Type A systems, but that personal-social characteristics were present and functioning as less dominant members. Had the latter been dominant, teacher success should have been associated with personal-social characteristics in systems of this type, which it was not.

In order to construct a paper and pencil instrument to assess the type and organization of norms appropriate to the project, three major steps were necessary, first, an appraisal of the universe of teacher behaviors and characteristics, second the reduction of this universe to a manageable set of items, and third, the organization of these items into a set of non-ipsative scales.

Three sources were employed in appraising the universe of teacher behaviors and characteristics which might be used in the instrument. First, items from a checklist of teacher behaviors of importance to success, developed under the preceding study. This list had been used in a survey of 70 principals involved in the study, who had checked the items they thought to be most important and added ones they thought to be important, but were missing from the list. Second, the universe of items appearing in the rating forms used by school systems in continental United States.

These forms were obtained through a stratified disproportionate random sample of school districts, with 50 percent of the systems above 50,000 population and 20 percent of the cities 20,000 - 50,000 being drawn. Third, the teacher characteristics measured in the project.

From this pool of characteristics four conceptual dimensions of teacher behavior-characteristics were derived. These dimensions were thought of as broad norms for teacher characteristics. The first was an organization norm and focused on behavior-characteristics aimed at classroom organization, management, and control. Although variously described, this norm appears in practically all rating scales used in school systems; it also appears among the characteristics used by Ryans. A second norm was labeled "learning". This norm embraces attention to individual differences, ability to diagnose learning difficulties, and the sequencing of learning materials. It is, of course, associated with teacher task performance characteristics, but it also occurs across rating scales, although its appearance was less frequent than were characteristics associated with an organization norm. A third dimension was labeled "emotional" and focused on characteristics associated with giving affection and showing emotional support. This norm appears directly on some rating scales, and it received some emphasis from principals, and was a distinct dimension in Ryan's original TCS, appearing as the attitude scales toward pupils, and toward democratic pupil practices. A fourth dimension was labeled "social" and encompasses emphasis on student social behavior. The latter received some emphasis by principals.

The second step in the construction of the TBWS consisted of writing items explicating each of the four dimensions. For the most part these items were drawn from the rating scales obtained from the survey of rating forms used in school systems, from the list checked by principals, and from Ryans' observation items underlying the scales of the TCS. The resulting pool of items were then placed on cards and individual faculty members and graduate students were asked to sort the cards into four homogeneous groups. Since the items were designed to be perfectly homogeneous within norms, any item which failed to be correctly sorted by a sorter was either discarded or revised. This process continued until four naive sorters ran perfect trials on all items. The items were then placed in a trial weighting instrument and administered to a group of graduate students. The point in constructing a trial instrument was to make certain that the items representing each norm were of the same order of generality.* The inference that an item differed in generality from the others was made on the frequency with which the item was given heavy weight or very light weight across raters. Items which yeilded a poor distribution of weights were eigher discarded or re-written, and the sorting process begun once more.

---

*There is also an implicit desirability dimension in each item, but this dimension is confounded with the broadness-narrowness or generality of the item.

The above procedures were continued until a point was reached at which five naive sorters sorted 12 items of equal generality into four homogenous dimensions. The twelve items were then divided into two groups of six each with one group of six labeled "task oriented," containing two norms, "learning" and "organization," and a second group of six labeled "social-emotionally oriented" with two subordinate norms, "emotional" and "social." These two sets of six items were then cross-paired 22 times in such a way that the three items under the organization norm recurred twelve times against the six items from the social-emotional set, and three items under the learning norm recurred ten times against the social-ecmotional set, and three items from social norm and the three items from emotional norm occured eleven times each against the members of the task oriented set. These pairs were then arrayed in an instrument which showed a weighting scale beside each member of each pair. The instructions to the respondent defined the meaning of each of the three weights beside each members, and instructed him to circle the number next to one member of each pair relative to the other member according to its weight as a factor in success in his school situation. A sample of pairs is shown below, and the full instrument appears in Appendix 2.

Is skillful in adapting learning tasks to individual differences	3	2	1	1	2	3	Readily shows affection for pupils
Encourages pupils to form small working groups	3	2	1	1	2	3	Is consistent and firm in managing pupils

The instrument may be scored for either four norms or two dimensions. To score on four norms, the weights of each homogenous set of three items are summed across the particular pairs in which they appear. To obtain a score for two dimensions, the weights of the learning and the organization norms are summed and the weights of the social and the emotional norms are summed.

While this instrument technically achieves the end toward which it was designed, that is, the ordering of sets of teacher behavior-characteristics which seem to be widely regarded as normative, by its very construction certain defects occured. For example, cleanness and neatness of appearance is virtually universal as an item on rating scales, and it is unquestionably utilized by principals and supervisors as a criterial characteristic or norm for teachers, If, however, one places this type of item with items like those shown in the sample, the clean-neat item will consistently appear with a light weight. Along either a dimension of generality or perhaps of desirability, the item is not competive with other items. Thus some norms for teacher behavior fail to appear in the scale because they are non-competitive in a rating scale of the type constructed, but are undoubtedly norms in actual school situations.

A second defect with the scale lies in the "hard choices" it requires. There is no mid-point of equal weight between the members of each pair, hence, no opportunity for equal weighting. Some respondents appeared to find this fact so frustrating that they avoided responding to the scale.

A component in the behavior modification-teacher selection model which is of equal if not greater importance than the norms for teacher behavior is that embracing community social and economic variables. While the community in which the school system is located is not the only source of norms for teacher behavior-characteristics,* it is clearly a proximate environmental source.

In the preceding study of beginning teachers, two community variables were associated with differences in norms apparently existing between school systems. The first was the per pupil wealth of the district (adjustedassessed valuation per resident pupil ADA), and the second, the ratio of middle class to working class schools. In the study, however, these two variables were confounded in such a way that cities high in per pupil wealth also had a disportionate number of schools with predominantly working class children, while cities low in per pupil wealth had a disportionate number of schools with predominantly middle class children. Thus, whether differences in norms between systems were associated with the prevailing socio-economic class of the community, or with wealth, could not be fully disentangled. There was evidence, however, that as the proportion of middle class schools increased, the association of beginning teacher success with personal-social characteristics grew increasingly sharp, as indicated by a rise in statistical significance levels. The latter result suggested that placing emphasis on the personal-social characteristics of the teacher was primarily a function the extent to which the community population was middle class. The subsequent study by Quinn (described in Chapter 2) shed further light on this question by showing that within a large urban community with a single school system, the personal-social characteristics of teachers were associated with success in the middle class schools, but not in inner-city schools. _In toto_ the pattern of results reviewed above suggested that two sets of variables might be at work, the first set comprised of variables associated with pupil wealth, the second set those associated with the socio-economic status of the community and possibly of the neighborhood.

The interpretation of the relationships of these two sets of variables to each other and to the components of the models under consideration, and therefore to differences between communites in teacher characteristics was of substantial importance in the project. As a variable, per pupil wealth was interpreted to be relatively independent of the central socio-economic characteristics of the communities. This assumption of independence arises primarily on the grounds that quite different sources of property wealth appear in communities of different types. Thus, in some communities a high level of wealth is based on upper middle class residential property, in some on large farm holdings, and in some on industrial-commercial property. From knowing only that a community has substantial property wealth, one cannot accurately infer the socio-economic characteristics of its population.

---

*Instruction in professional schools is an alternate source.

Beginning with the assumption of independence between per pupil wealth and the socio-economic characteristics of communities, the further inference was made that if differences between communities in per pupil wealth is in some way associated with differences between them in teacher characteristics, the latter association could arise either because wealth in some way offered a special attractiveness in communities for elementary teachers or else wealth operated largely as the principal economic variable determining the kind of characteristics displayed by the school system. In the case of the behavior modification-teacher selection model, only the latter possibility was available. Observing the latter as the only apparent possibility, and further observing that recent work by Kiesling (12), and by James and others (11), indicate that per pupil expenditure is in significant part, but not wholly, associated with per pupil wealth (per capita wealth was used by Kiesling), the best available inference appeared to be that per pupil wealth operates through per pupil expenditure, and that the systemic characteristics associated with the latter would be important in the study.

In the James study, two characteristics associated with expenditure, and relevant to the present project, were pupil-teacher ratio and median starting salary. In an earlier study, Kiesling also found these systemic characteristics together with teacher salary incentive, to be associated with expenditure. The difficulty with these particular characteristics is that they do not neatly fit the behavior modification model, rather, they are attraction variables, and therefore are best oredered to the differential attraction model. As noted earlier, a variable which the author found to be associated with wealth-expenditure, and relevant to the behavior modification model, is the presence of a central supervisory staff in the school system. This variable is not a simple function of wealth or expenditure, but appears as a joint function of wealth and system size. Small wealthy systems no doubt find it difficult to justify a full time supervisory staff.

At this point it should be noted that the identification of per pupil wealth as a community characteristic does not lead, under the evidence available, to variables which could be rationally associated with systemic norms. Rather, it leads to teacher attraction variables on the one hand and in conjunction with size, to what was earlier labeled a systemic "delivery mechanism" of the other. Thus, if per pupil wealth is retained as community characteristic relevant to differences in teacher characteristics, a subordinate model is implied: namely

Per pupil wealth - - - - - - - - - - - Supervision (delivery mechanism) - - - - - - - - - Differences in teacher characteristics

If a little thought is given to this model in relation to the over-all behavior modification-teacher selection model of which it is a part, it becomes apparent that the modification of teacher characteristics and/or the differential selection of teachers after they are employed requires the operation of two orthogonal sets of variables, namely, wealth, and socio-economic status. Socio-economic status would be hypothesized to be the determinant of the norm toward which behavior is modified, while wealth, probably in conjunction with size, would be the determinant of whether the means or mechanism to modify the behavior is available.

Since the relationship between socio-economic status and the norms apparently used by school systems were earlier elaborated, with task orientation associated with proportionally more working class schools in the system, and a personal-social orientation associated with middle class schools, the central remaining question was which particular variables would be chosen to represent class status. The four variables chosen were median family income, median education level for adults 25 years and older, number of college graduates per thousand, and percentage of population engaged in manufacturing. Of these variables, the first three were regarded as direct socio-economic measures, the fourth variable, however, is not a direct measure, but rather an estimate of the nature of the working class population, i.e. whether it is an industrial working class.

A subordinate variable taken in conjunction with socio-economic status of communities was the predominant socio-economic status of specific schools within the community. The point of using this variable was to isolate the extent to which norms generally attributable to community S-E level might be localized. Thus, although a teacher might be employed in a community below average in family income, education level and college graduates per thousand, she could still be employed in a middle class or upper middle class school within the community. To the extent that norms are both class and neighborhood bound, her characteristics, or the value placed on them, might be anticipated to deviate from teachers in the same community employed in working class schools.

A final consideration to be taken into account in the behavior-modification-teacher selection model lies in the teacher characteristics used in the appraisal. The assumption that a school system modifies teacher behavior or characteristics clearly involves the assumption that the behaviors observed are modifiable. This assumption is clearly justified with respect to task performance, child versus-subject centeredness, and possibly justified with respect to Warmth-spontaneity and Stability. It follows that to the extent that systems might modify teachers, differences in levels between the more modifiable and the less modifiable characteristics might be anticipated to occur.

Systemic Characteristics and the Differential Attraction Model. This model assumes the same social and economic characteristics as the modification-selection model, but involves the alternate hypothesis that these characteristics lead to the differential attraction of teachers in addition to, if not rather than, the modification of teacher behavior. As suggested in the diagram of the model presented earlier, the attraction of teachers to a system may be viewed as a function of either particular policy features of the school system which are specifically attractive, i.e. salary and working conditions, or of features of the community. Interestingly, this dichotomy between whether it is the system or the community which might be the greater determinant is associated with the separation of per pupil wealth from the socio-economic characteristics of the community. The association occurs primarily on the grounds that to the extent per pupil wealth is a determinantof per pupil expenditure, so is it a determinant of the particular features displayed by the school system. Wealth and expenditure are not only associated with increased availability of supervision, but also with salary limits, salary incentives, pupil teacher ratio, and other features of systems which might be anticipated to be attractive to teachers, especially to teachers for whom teaching is primary rather than secondary employment.

For most elementary teachers, however, teaching might be regarded as secondary employment. Most are female, and most of the females are married. The primary employment in the household of the typical married female teacher would be the occupation of the husband, while the occupation of the wife would be secondary. This particular economic arrangement in the households of elementary teachers is probably of great significance in the present study, since it casts the attractiveness of the community for the employment of the husband as the primary consideration, with the attractiveness of the community or the school system to the wife, as a teacher, following as a secondary consideration. In essence, the pool of potential elementary school teachers for a system may be viewed as a function of the attractiveness of the community for the husband. The husbands of elementary school teachers, in turn, may be anticipated to be husbands of a particular kind, namely college graduates, and professional or managerial or perhaps technical and kindred workers (13).

Following these notions, it appears to be the case that the attractiveness of a community for the husband would be a function 1) of the proximity of employment in industrial or commercial firms, or employment in providing professional services to the employees of such firms, and 2) of the wage levels of the community, which might be indexed by median family or per capita income. In addition to the two factors above, there are two, and possible three other factors to be taken into account relative to the husband and wife, and incidently, to the single female teacher. One of these is the educational level of the community, the second, college graduates per thousand, and the third, community size.

Community education level and college graduates per thousand are variables closely colinear with median family income. Together, these three variables provide a reasonable index of the general cultural level of the community. The cultural level of the community may in turn be viewed as an attraction to the female teacher, married or single, and also perhaps as an attraction to the husband of the married teacher.

The entry of community size into the matrix of variables under consideration occurs primarily in conjunction with Standard Metropolitan Statistical Area (SMSA) core city and suburban areas. Typically, the core city area would be expected to contain the industrial and commercial property wealth, and also the employment of the husband as a primary wage earner. Thus, one might say that the core city attracts the employment of the husband. The suburban areas, however, hold the status variables, and attract the residence of the husband and wife, and at the same time offer to the wife secondary employment under culturally advantageous conditions, if not also under specific systemic conditions that are also attractive, i.e. high achieving pupils (11). Given a choice between the core city system and the suburban system, the married female elementary teacher might indeed opt for the suburban system, thus giving it some advantage in the marketplace of teacher characteristics.

Under the differential attraction model it is of course hypothesized that some communities are in a better position to attract teachers than others. Such attraction would operate to produce differences between systems, however, only under one of two conditions. First, that all teachers in the potential employment pool were very homogeneous, and the mean characteristics of the pool were different from those in the pool available to systems with less attractiveness. In this instance, the advantaged system could draw at random from the pool and obtain a group of teachers different from disadvantaged systems. The second condition is that the pool available to advantaged systems is very like the pool available to disadvantaged systems, but is proportionally larger, i.e. the supply of teachers relative to demand is greater. In this instance, if the advantaged systems differentially selected from the pool, differences between systems would arise, if the advantaged system could predict upon employing a teacher whether she would be congruent in her behavior with the norms in the system. Between the two, the second was given greater weight in the project, primarily on the assumption that when a school system can selectively employ teachers it moves to do so whether or not the basis of selection accurately predicts the subsequent congruence between the behavior of the teacher and the norms of the system.

A final problem associated with the use of alternate models through which differences in teacher characteristics between systems occur, lies in ascertaining which type of model might be operating. In a cross-sectional study such as the present project a final solution to this problem cannot be achieved; however, by examining shifts in teacher characteristics according to the years of experience a teacher has in the system in which she is employed, relative to other systems, relatively firm inferences can be made. For example, if teachers with 0-5 years of experience in their system show homogeneous characteristics across systems, while teachers with six or more years of experience in their system show differentiated levels of characteristics, evidence for the differential attraction model is weak and for the modification model strong. On the other hand, if there are differences in the characteristics of teachers 0-5 years experience in their system across systems, and these differences are maintained in the group with six or more years of experience in the system, evidence for the differential attraction model is strong, and the modification model weak. There are, of course, many variables both in school systems and in particular schools which can function to obscure clear-cut relationships between systems. Many of these variables can be controlled, however, and the next chapter discusses the means by which some of these variables were identified and brings into final focus a problem left unresolved in this chapter, the empirical relationships among school system and community characteristics and the generation of school system types.

CHAPTER 4

SAMPLE SURVEY PROCEDURES AND THE GENERATION OF SCHOOL SYSTEM TYPES

Sampling Design. A three stage stratified cluster quota design was employed in the study. The first stage units were school systems, for which the population was the 97 Indiana school systems with enrollments above 1800 students, grade 1-12 in 1964-65. From the population 19 systems were removed by reason of having 1) participated in previous research utilizing similar instruments, 2) employed a new superintendent within 18 months, or 3) having been listed as having departmentalized instruction in grades five and six. The remaining 78 systems were stratified by enrollment grades 1-12 (3 strata), by per pupil wealth (2 strata), and by tuition tax rate adjusted to equalized valuation (3 strata), with the latter variable functioning as a control on education effort. In addition, the location of each district in one of three state regions (north, central, south) was recorded. Sampling was subsequently conducted on a quota basis, drawing from within each enrollment-wealth cell by tax rate and by region.

The second stage clusters were schools within systems, with the number of schools to be drawn proportionate to the number of schools in the district stratified according to the socio-economic class of the neighborhood, (middle class, mixed, working class) as reported by the superintendent. A limit of six schools was placed on larger systems, since the project staff could test in no more than six schools in a single day. The possibility of testing on two separate days was ruled out in order to eliminate the possibility of communication between teachers in the same system.

In the third stage of sampling, two sets of elements were drawn, 1) the principal of each school, and 2) all of the teachers in each school, excluding special teachers. It should be noted that the desirability of drawing all elements within the second stage clusters rested in part on the dispersion of the second stage clusters across S-E strata, so that in the event there was homogeneity of teacher characteristics within schools, but differences between schools, the estimate of the characteristics of teachers within systems would still be based on heterogeneous sample within systems, i.e. the first stage clusters.

The central advantage to this design was that it was relatively amenable to adjustment as various unexpected contingencies developed without excessive loss of representativeness or external validity. Random sampling cannot be claimed, of course, but high non-participation rates at the first stage level effectively preclude true randomness in any event, and the point of the design is maximize the probability of external validity, hence replicability, when it is known before hand that random selection cannot be effectively attained.

The central difficulties with the design came at two points. First, less than half (20/42) of the systems approached were willing to commit teachers to approximately 2.5 hours of participation required in the project. This problem was anticipated, as decribed in Appendix 1, but its magnitude was greater than anticipated. To adjust to it, the enrollment-wealth cells were maintained, but within cells, region was compromised first, and tax rate second. As may later be observed in Table 4-1,(page 43) there are relatively few systems in the sample from southern Indiana, and those present are small. On the other hand, the population centers in Indiana lie in the central and northern sections, and the over-all effect on the estimates of teacher characteristics for the state is perhaps not severe.

The second point at which difficulty was experienced lay in maintaining the proportionality in both the second stage clusters and in the number of teachers entering the study from particular systems. These difficulties developed out of a complex set of relationships between the number of schools the superintendent was willing to involve, the dispersion of these schools over S-E strata, and variablity in the number of teachers in particular schools. Of the three factors immediately above, greatest emphasis was given to maintaining an appropriate dispersion of schools over S-E strata within systems.* Thus, in small systems with distinct neighborhoods, a disproportionate number of schools were drawn, if the superintendent was willing to involve several schools, which he usually was. In large systems superintendents were typically willing to sample from each S-E stratum, but were unwilling to involve a large number of schools. Since the schools sampled tended to be larger in large systems than in small systems, the effect of disproportional numbers of schools in creating a disproportional number of teachers from certain systems was less drastic than expected. As may be observed in Table 4-1, however, there is some tendency for the smaller systems to be over-represented in terms of the number of teachers from them appearing in the sample.

---

*Emphasis was placed on maintaining a dispersion of schools over S-E strata in order to maintain correspondence between the principal S-E indicators for the community, i.e. median education level and median family income, and the number of schools drawn from each type of neighborhood stratified by S-E. For example, a community below average in both median income and education level probably cannot be represented by a middle class school. The relationships between the community S-E indicators and the S-E stratum from which particular schools were drawn within each system may be observed in Table 4-1. It should be noted that "mixed" schools (category 2, Table 4-1) contain both working class and middle class children, and are typical of many smaller systems in which there are no large, distinct neighborhoods with contrasting S-E status.

In sampling teachers within schools few difficulties developed. Participation rates were in excess of 90 percent in completion of the problem-tasks, and about 75 percent in completion of the Characteristics Schedule. As may be observed in Table 4-1, however, a disproportional number of primary teachers appeared in the sample. This effect arose in part from a slightly better participation rate among primary teachers, and in part from the way in which instruction was organized in certain systems, with several schools having an excessive number of primary classes and fewer intermediate classes. In such instances the intermediate pupils were either sent to a departmentalized intermediate school after grade five, or simply moved to another school, and their teachers did not appear in the sample.

Procedures in School Systems. The superintendents of all systems approached in the study were first contacted by telephone, and the purposes of the study, and the general sampling plan for obtaining schools within the system, described. A letter detailing the study was then sent to each superintendent, who subsequently discussed the project with the principals of the schools tentatively to be involved. A firm agreement concerning whether the system would or would not participate was typically reached at this point. If the system decided to participate, the investigator paid a visit to it to meet with the superintendent or his designate (the director of elementary education, the curriculum coordinator or an assistant superintendent) and the principals. During this meeting copies of the instruments to be employed with teachers were inspected by those present, and confidentiality of the specific content of the instruments requested by the investigator. In addition, the investigator explained that teachers in each participating school would receive a letter from the project center inviting them to participate in the project in their school building after school for about one hour and fifteen minutes, that they would be asked to complete a preference schedule at home for an additional hour, and that confidentiality was assured. Principals were informed that they would be interviewed for approximately one hour by a member of the project staff on the day the teachers were tested, and that they would subsequently be asked to complete the TBWS and to sort a deck of cards, each card bearing the name of one of their teachers, into four equal groups according to over-all success in teaching in that specific school. All present were assured that both the scores of individual teachers and the ratings by the principal would be held in the strictest confidence, but that a report to each teacher would contain his individual scores and an interpretation of the project, while a special report sent to administrators would give the average scores for the system in relation to other similar systems in the state and in relation to the norms for the state.

TABLE 4-1. Number of Schools and Teachers Sampled in Relation to Selected Community Characteristics.

System	Enrollment 1964-65	Median Years Educ. 1960	Median Family Income	Number of Schools by S-E status*			Number of Teachers Tested		Location in State
				1	2	3	Primary	Interm.	
01	2,956	11.8	$4,830	0	2	1	16	11	C
02	2,074	9.9	5,045	0	0	1	10	5	S
03	1,819	10.9	6,425	1	1	1	13	14	S
04	3,110	11.9	5,996	0	1	1	15	13	C
05	8,502	11.3	6,980	1	1	1	21	20	N
06	5,651	9.9	5,187	1	0	1	11	9	C
07	44,289	10.0	6,004	0	0	3	37	28	N
08	2,048	10.3	5,106	0	0	1	4	3	C
09	5,160	11.7	6,910	2	0	0	18	17	N
10	7,179	11.2	6,156	1	1	2	27	24	N
11	10,789	10.6	5,890	1	1	1	35	19	C
12	17,514	10.5	5,607	1	0	2	23	18	C
13	5,239	11.1	6,555	4	3	0	31	29	N
14	2,848	12.5	7,574	1	0	0	10	8	C
15	2,623	10.6	5,771	1	0	0	6	7	C
16	20,824	10.7	5,292	1	2	2	28	29	C
17	3,035	11.0	5,818	0	1	1	14	11	N
18	2,628	10.0	4,808	0	1	0	5	6	S
19	9,222	11.4	6,300	1	0	1	14	11	C
20	2,125	10.6	5,781	1	1	1	17	18	N

*1=parents are professional or managerial, 3=parents are skilled, semi-skilled, or unemployed, 2=parents are drawn from all occupational categories, (as perceived by superintendent and principal).

The procedures outlined above were followed quite rigorously in all systems. Teachers in each system were met at the designated time, typically at 3:30 p.m., in a classroom in their school by a member of the project staff, who re-explained the purpose of the project and re-assured teachers that individual results would not be revealed to any persons or agencies outside the project staff. The problem-task booklets and a one page data sheet* were than handed out, and each person instructed to read the instructions to each task carefully, to proceed at his own pace, and to raise his hand for help if he was not perfectly clear about what to do in performing a task. As each teacher finished, the staff member collected his materials and showed him the Characteristics Schedule and an answer sheet, and explained to the teacher how to record his responses. Both the Schedule and the sheet were then enveloped and the participant instructed to return them to the project center at his earliest convenience.

Both the TBWS and the rating were forced choice, and ultimately three of the principals said that all of their teachers were superior (in two of the three instances, a judgement subsequently borne out by the mean levels of the characteristics of their teachers), and that they could not rate them on a forced choice basis, while five principals failed to return the TBWS, apparently by virtue of inability to make the choices required.

The Generation of School System Types

The purpose of generating school system types was to provide a useful scheme for summarizing patterns of school system and community characteristics. The utility of a particular scheme depends in part upon what one can predict by using it, and in part on its replicability or generality. An ad hoc set of types which can in selected instances be used to make predictions, but have little generality beyond the sample in which they are employed, are not useful, but neither is a highly general or easily replicable set of types which hold only weak associations to criterion variables of interest. The problem in the project was to isolate a set of types which fulfilled both of these utilitarian criteria.

To establish coherent patterns of school system and community characteristics, the principal independent variables discussed in the preceding chapter, together with adjusted tuition tax rate were factor analyzed by the principal components method with varimax rotation to orthogonal factors, with subsequent generation of factor scores. The observed values of the variables for each system were drawn from several different sources. Median family income, median education level (adults 25 years and over), percent population in manufacturing, and college graduates per thousand population were drawn from data in the 1960 census.

---

*See Appendix 2.

Enrollment, per pupil wealth, pupil teacher ratio, and pupil cost were drawn from the Indiana Department of Public Instruction, Statistical Summary, 1964-65.* Terminal salary with the M.A. and salary incentive (difference between beginning salary with the M.A. and Terminal salary with the M.A.) were drawn from Circular No. 10, 1966, Indiana State Teachers Association. Systemic values for enrollment, median family income and median education level may be observed in Table 4-1, and for the remaining variables in Table 4-2. The three factors extracted and the loading of each variable on each factor may be observed in Table 4-3.

In interpreting the factors, emphasis was placed on "source" variables as well as factor loadings. In factor 1, per pupil wealth is the apparent source variable, from which flow pupil expenditure, and variables contingent upon pupil expenditure, i.e. P-T ratio, and salary limit and incentive. This factor was therefore labeled "wealth-cost," but it contains those variables which clearly serve as employment attractions to teachers.

The second factor, urbanization, was interpreted to have two source variables, enrollment and percent of population engaged in manufacturing. In the preceding study this dimension was confounded with wealth. In the present study its appearance separately from wealth was in part contingent on entering percent of population engaged in manufacturing as a variable, and part contingent upon sampling districts in which wealth was not tied to industrial sources, but to rural property, thus dispersing the wealth variable across non-industrial communities. The high positive loading of tuition tax rate on this factor seems also to follow from this urban-rural split. Urban systems rather uniformly have a higher tax rate, independent of their wealth, and apparently independent of their median income levels. Why this is the case is not altogether clear, but a complex interaction between educational demand, wealth from industrial-commerical sources, and median family income levels appears to be the undergirding factor. In this respect it may noted that in urban-industrial districts, either industrial-commerical property wealth or high median family income may offset the weight of the tax burden on individuals, but in rural communities the tax burden appears to rest largely on individuals, not corporations, and the individuals involved tend to fall on the low side with respect to median family income. In addition there is, perhaps, less educational demand in the rural community.

---

*Data from 1965-66 was not yet available when sampling procedures were started.

TABLE 4-2. School and Community Social and Economic Variables.

System	Pupil wealth	Adj. tution tax rate	Pupil cost	T-P ratio	Teacher salary incentive	Teacher salary limit	% pop. in mfring.	Coll. grads. 1,000 pop.
01	$15,742	2.59	$512	18.8	$4,225	$ 9,625	14.6	56
02	2,844	3.34	389	30.6	3,100	8,600	33.2	62
03	9,084	2.48	476	29.2	3,520	9,420	32.8	45
04	8,349	2.22	370	28.3	2,600	8,000	32.5	78
05	10,217	3.33	511	21.9	4,350	10,170	44.5	77
06	6,899	2.71	418	25.9	3,600	9,500	47.9	41
07	9,852	5.62	519	25.7	4,400	10,800	50.6	51
08	10,690	2.75	428	33.1	3,300	9,100	25.0	56
09	5,672	5.13	406	25.4	3,600	9,800	50.6	63
10	9,647	3.60	447	23.0	4,256	10,192	41.0	69
11	8,914	3.33	478	24.7	3,930	9,804	43.2	59
12	7,069	3.38	417	25.3	4,100	10,000	40.0	75
13	8,570	3.97	472	25.7	4,200	10,248	41.0	80
14	7,009	3.66	344	30.0	3,600	9,300	29.2	142
15	9,664	3.19	475	25.8	3,700	9,600	35.0	76
16	9,254	3.62	530	23.8	3,825	9,435	26.1	67
17	7,795	3.52	422	20.4	3,400	9,250	44.5	58
18	8,080	3.81	432	22.7	4,140	9,440	16.8	56
19	8,392	2.73	540	22.9	4,300	10,000	31.4	92
20	7,919	2.84	390	23.6	3,100	8,900	30.8	39

TABLE 4-3. Factors Extracted from 11 School and Community Social and Economic Variables (Principle components, varimax rotation to orthogonal factors).

Variable	Factor 1 Wealth-cost	Factor 2 Urbanization	Factor 3 Income-education
District enrollment	.35	.70	-.21
Per pupil property wealth	.80	-.36	.02
Adjusted tax rate	.07	.84	.04
Per pupil cost (expenditure)	.85	.17	-.11
Pupil-teacher ratio	.70	.01	-.02
Teacher salary incentive	.82	.43	.11
Teacher salary limit	.66	.68	.11
Median family income	.13	.38	.84
% population engaged in manufacturing	.20	.80	.10
Median education level	.09	-.25	.89
College graduates per 1,000 population	.03	.01	.82
Eigenvalues	3.75	2.38	2.13
Cumulative proportion of variance	.34	.56	.75

The third factor, income-education, is highly unitary and may be construed as representing community socio-economic status. The extraction of this factor as separate from both wealth and urbanization rests in substantial part on the appearance in the sample of several districts (04, 09, 13, 14, and 17 are most prominent), which have high median income and high median education levels, but were in the middle to low range in enrollment and in property wealth, and dispersed over the range in percentage of population in manufacturing.

The extraction of three factors from the 11 school system and community variables originally put into the matrix provided three major alternatives for the generation of school system types. The first alternative was the generation of eight types derived by dichotomizing the factor scores for systems on each of the three dimensions extracted. There were two difficulties with this approach. The first was that there was no guarantee that the types thus generated could be replicated, since the factor scores were recognizably contingent upon the intercorrelations among the variables in the particular sample of 20 systems. The second difficulty lay in dividing 20 systems eight ways, which led to too few observations in some cells whether one counted systems, schools or teachers as degrees of freedom.

The second alternative was to utilize the three factors as guides in producing an intuitive set of types. This approach is intuitive because the production of four or five types of systems from the three factors requires some conceptual method for combining the factors. A substantial amount of time and analysis were devoted to creating a set of types under this alternative. The most workable set obtained was generated on the assumption that the very large urban systems (SMSA core cities, enrollments in excess of 15,000) faced a different set of circumstances than other systems in the study. The three urban systems were 07, 12, and 16, each of which also had below average median education levels. The remaining systems were then ordered to three types in such a way that there was a generally descending order on both the urbanization dimension and the income-education dimension. The attractiveness of this set of types lay in the fact that the systems in two of the four types were observably similar. Thus, the large urban systems were clearly large, urban, and industrial, while the non-urban, low income-education level systems were clearly identified with small, rather sleepy Indiana towns. There were, however, several difficulties with these types. First, they were very heterogeneous on the wealth dimension. Second, unless the wealth factor was introduced in an ad hoc way, there were no clear lines for dividing systems in the middle range, i.e. systems between the large, urban communities and the sleepy towns. The distinction in the middle range was made on the basis of factor 3, (income-education) but the cutting points on the variables were very fine, thus probable error in the grouping was high, and the probability of replication low. Finally, this set of types clearly confounded major variables in an unnecessary way in terms of the models discussed in the preceeding chapter, and it was therefore ultimately abandoned.

The third alternative in generating systemic types was the utilization of wealth and income-education as the dominant variables, disregarding urbanization as a factor. There are two advantages to a set of types thus generated. First, by representing each dimension through a single variable, per pupil wealth in the first instance, and median family income in the second, a highly replicable set of systemic types were generated, since these variables are almost universally available. Second, the use of the latter two variables made possible a direct connection between the models discussed in the preceding chapter, and the systemic types. The disadvantage of these types was primarily empirical and stemmed from the appearance of an urbanization dimension of communities. The extent to which urbanization is actually disruptive to the types depends, of course, on the association between urbanization and teacher characteristics. A matter obviously open to empirical investigation. Conceptually, however, there is some disruption in placing together in the same type or cell systems which, on the surface, are very different, for example, a farming community in which there is no village in excess of 1,500 persons and an SMSA core city in excess of 70,000 persons. With respect to teacher characteristics, however, there was no explicit reason to suppose that urbanization, independent of wealth or income-education, is a determining factor in teacher characteristics and in such instances it seemed the more judicious course to entertain the null hypothesis until the data required its rejection.

To obtain the specific types of system used under the alternative described above, per pupil wealth was dichotomized at the sample median, which was closely approximate to the state median, and median family income was also dichotomized at the sample median, which again was virtually identical to the median family income in the state. The four resulting types were labeled as follows:

Type 1--hi wealth, hi median income	(6 systems)
Type 2--lo wealth, hi median income	(5 systems)
Type 3--hi wealth, lo median income	(4 systems)
Type 4--lo wealth, lo median income	(5 systems)

After the systemic types were established, the procedures in the study became specific to the particular type of statistical analysis employed. For the analysis, all data from teachers and principals were coded, as shown by the coding sheets in Appendix 2, and placed on IBM cards. A CDC 3400-3600 computer system was used for data processing. Since the statistical models used in the analysis varied according to the question to be answered or the hypothesis to be tested, the attendent procedural details are presented in the following chapter together with the analysis pertinant to the particular point under consideration.

CHAPTER 5

PROCEDURAL DETAILS AND RESULTS

As noted earlier, there were two problem-tasks in teaching reading and two in teaching arithmetic. Since each of these tasks had differing raw score totals, the raw scores for each were transformed to Z scores, with a mean of 50 and a standard deviation of 10. These Z scores were then summed, producing a mean Z score of 100 in reading and a mean Z score of 100 in arithmetic. The scores on the problem task in science and the scales from the Characteristics Schedule were also transformed to Z scores with a mean of 50 and an S.D. of 10. In order to take differences in the problem-task and differences in populations into account, the Z score transformations were done separately for primary and for intermediate teachers. Among the problem tasks, there were differences in the raw score - Z score correspondences for primary and intermediate teachers; among the scales of the Characteristics Schedule there were no such differences, indicating that the score distributions of primary and intermediate teachers on these scales were functionally identical.

Relationships Among the Tasks and Scales. For the first analysis, teachers who taught special areas in either the primary or intermediate grades and who therefore did not complete all the problem-tasks, teachers who did not return the Characteristics Schedule and teachers with less than two year's experience were eliminated from the sample. The task and scale scores of the remaining group of primary and intermediate teachers were then factor analyzed separately using a principal components method with varimax rotation to orthogonal factors. The resulting correlation matrices may be observed in Table 5-1 and the factors and factor loadings in Table 5-2.

As shown in Table 5-2, the factor structures of the tasks and scales are highly similar for primary and intermediate teachers. The first factor extracted for each of these groups is clearly identified with the scales of the Characteristics Schedule, which consistently show moderate correlations with each other. Among these scales, only Viewpoint differentiates to load on factor 2, a factor associated with task performance in teaching reading and arithmetic. The negative sign of the loading of the Viewpoint scale on factor 2 indicates that higher performance on the reading and arithmetic tasks is associated with lower or more child-centered scores on the Viewpoint scale. As may be noted in Table 5-1, however, the magnitude of the correlations between the Viewpoint scale and the tasks is very modest indeed and the relationship might be interpreted as present but very weak.

The third factor shown in Table 5-2 is largely associated with the science task and to a lesser degree with the Induction scale, which holds a very small positive relationship to the science task among primary teachers but a somewhat stronger relationship among intermediate teachers. The low reliability of the Induction scale, together with its failure to

TABLE 5-1. Correlations Among Task and Scale Scores for Primary (above the diagonal, N=282) and for Intermediate Teachers (below the diagonal, N=214).

	Tasks					Scales			
	Read.	Arith.	Sci.	W-S	Org.	Vpt.	Stab.	Involve.	Induct.
Read.		21^{xx}	11	20^{xx}	08	16^{xx}	.07	-06	09
Arith.	34^{xx}		02	00	05	-08	-08	-01	03
Sci.	13	18^{xx}		09	09	05	03	-04	08
W-S	12	12	12		37^{xx}	-42^{xx}	50^{xx}	20^{xx}	11^{x}
Org.	00	09	10	45^{xx}		-07	48^{xx}	24^{xx}	10
Vpt.	10	12	01	-47^{xx}	-11		26^{xx}	12^{x}	03
Stab.	03	09	05	62^{xx}	50^{xx}	-44^{xx}		29^{xx}	09
Involve.	05	03	03	32^{xx}	40^{xx}	-14^{x}	34^{xx}		08
Induct.	16^{x}	16^{x}	28^{xx}	17^{x}	19^{x}	-07	21^{xx}	11	

x, p=05
xx, p=.01

TABLE 5-2. Factors and Factor Loadings from Task and Scale Correlations, for Primary and for Intermediate Teachers.

	Primary Teachers Factors			Intermediate Teachers Factors		
Variable	1	2	3	1	2	3
Read.	03	67	28	02	74	13
Arith.	-17	66	02	03	74	15
Sci.	-02	08	74	02	19	71
W-S	73	34	09	81	22	02
Org.	67	-11	30	71	-17	33
Vpt.	-45	-55	20	-56	-44	41
Stab.	81	01	07	90	09	00
Involve.	56	-16	-11	59	-16	21
Induct.	11	01	60	21	20	64
Eigen Values	2.32	1.34	1.04	2.73	1.48	1.11
Cumulative % variance	26	41	52	30	47	59

hold more significant relationships to the problem-tasks (a fact probably associated with its low reliability) largely eliminated its utility in the study, and it was dropped from the analysis after several preliminary runs made apparent that it was not closely linked to any of the dependent or any of the independent variables.

Relationships of Task and Scale Scores to Major Community Social and Economic Variables, with each School District as one Observation. The relationships between teacher performance and community variables may be examined both under the concept that each school system represents one observation or degree of freedom, and under the concept that each teacher represents one observation or degree of freedom. When each school system or district is taken as the unit of observation, the observed scores of the teachers in that system are represented by their means. Such means may be viewed as unweighted in the sense that they do not directly reflect differences between systems in the number of teachers actually undergirding the mean values. Each system contributes equally as it were. The central disadvantage with the use of these means in the project lay in the fact that some very small systems were represented in the primary grades by as few as four teachers, which introduced the possibility of unreliable extimates of true mean values, potentially contributing to the error variance between systems within the same type.

When each teacher is used as the unit of observation, the situation is approximately the reverse of that described above. The mean values of each system enter the analysis in such a way that they are proportionate to the number of teachers drawn from each system. In this instance, the larger systems carry a heavier weight in the analysis than do the smaller systems, although this weight is of course proportionated to the number of teachers actually represented in these systems. The central disadvantage to the use of weighted means lies in the possibility that ad hoc variables operate in particular systems. If such variables operate in large systems contributing many teachers to an analysis, between system error variance may again be produced, or if the system is highly dominate in the cell to which it is ordered, error variance between system types may be produced, which when confounded with true variance between system types may lead to erroneous inferences.

Since somewhat different types of error are associated with the unweighted means analysis and the weighted means analysis, both forms of analysis were used in the project wherever feasible. Because only 20 systems were involved, however, there were rather severe limitations on the use of the unweighted means analysis, and the weighted means analysis became the preferred form.

For the first unweighted means analysis, school districts were dichotomized on median family income, then re-dichotomized on per pupil wealth, producing a 2 x 2 analysis of variance arrangement. To obtain the systemic means for teachers on each dependent variable, all classroom teachers, whether or not they taught in a departmentalized arrangement, were used[*], but teachers of special subjects and teachers who had entered the system within the past 18-24 months were excluded. These means were taken as the best estimates of the performance and characteristics of the stable core of classroom teachers in each system. The effects for primary and intermediate teachers were calculated separately, and may be observed in the upper and lower portions, respectively, of Table 5-3.

As may be observed in Table 5-3, there are no significant relationships between median family income or wealth and the task performances or characteristics of primary teachers. Among intermediate teachers, however, median family income is associated with each area of task performance, as well as with Warmth-spontaneity. Significant interactions also occur between median family income and per pupil wealth for Organization and for Stability. It may be noted that the means for the lo MFI - lo wealth districts typically lag those in other districts, distinctly contributing the main effects present, and contributing to a significant interaction in those instances in which the scores of teachers in lo MFI - hi wealth districts are disproportionately high.

To examine the relationships between urbanization as a dimension of communities and to cross-check relationships among variables associated with the wealth and socio-economic dimensions, all in relation to teacher task performance and characteristics, several additional analysis were conducted. The first analysis of this series was performed by dichotomizing districts according to size (enrollment), then re-dichotomizing by per pupil wealth relative to size, with factoral analysis of variance on each dependent variable. There were no significant main effects or interactions among either primary or intermediate teachers in this analysis. For a second analysis, districts were dichotomized by median education level and re-dichotomized by per pupil wealth, producing an analysis directly compariable to the MFI by wealth analysis, but with the contribution of the socio-economic dimension of communities estimated from median education level (MEL) rather than median family income. The results of this analysis were slightly different than for the MFI by wealth analysis. Among intermediate teachers, the main effects for median education level with respect to arithmetic (F=11.90, p .01) were weaker, but the main effects for science (F=11.42, p .01) slightly stronger than the comparable effects for MFI, there were, however, no significant main effects for Warmth-spontaneity and no significant interaction for Organization or Stability.

---

[*] The use of teachers in departmentialized schools precluded the use of the combined Z score for reading, arithmetic and science in this analysis, since departmentalized teachers performed only those tasks directly related to their teaching area, e.g. arithmetic.

TABLE 5-3. ANOVA for Task and Scale Scores of Primary and Intermediate Teachers, Per Pupil Wealth by Median Family Income, with School System Means as Observations.

	Cell Means						Effects		
	Hi MFI Hi Wealth	Hi MFI Lo Wealth	Lo MFI Hi Wealth	Lo MFI Lo Wealth	MSE	df	MFI F	Wealth F	MXW F
PRIMARY									
Read.	100.78	98.14	100.76	97.55	27.42	16	0.02	1.53	0.01
Arith.	98.85	99.47	102.83	101.49	38.48	16	1.14	0.16	0.12
Sci.	49.46	49.77	49.15	51.28	6.38	16	0.27	1.14	0.63
W-S	50.10	50.30	49.03	48.20	11.14	16	1.11	0.04	0.11
Org.	49.80	50.76	49.98	48.18	11.68	16	0.61	0.07	0.79
Vpt.	49.56	49.98	52.43	51.24	12.06	16	1.72	0.06	0.26
Stab.	49.92	48.90	49.15	47.66	19.02	16	0.26	0.04	0.01
Involve.	50.25	50.50	46.52	48.80	14.92	16	2.14	0.52	0.34
INTERMEDIATE									
Reading	102.36	103.64	100.29	91.77	29.48	16	8.07^{x}	2.18	3.98
Arith.	102.18	104.01	93.67	96.88	13.02	16	22.99^{xx}	2.39	0.18
Sci.	50.92	52.84	49.04	47.57	7.74	16	8.00^{x}	0.03	1.81
W-S	50.51	51.14	49.55	45.10	12.58	16	4.70^{x}	1.38	2.45
Org.	49.00	50.56	53.03	44.38	19.31	16	0.29	3.18	6.60^{x}
Vpt.	48.95	49.00	50.93	49.48	11.64	16	0.63	0.21	0.24
Stab.	50.00	50.30	52.85	47.62	7.35	16	0.00	4.16	4.97^{x}
Involve.	48.81	50.40	50.70	47.64	7.64	16	0.31	0.14	2.73

x, p=05, xx, p=.01

On the other hand, a main effect (F=4.15, p .05) and a significant interaction, (F=8.14, p 05) appeared for reading. Among primary teachers, the only significant effect was for wealth with respect to the Involvement scale (F=5.88, p .05).

The differences in effects between MFI and MEL are largely traceable to the interchange of three high wealth districts between the two analysis. In Table 4-2 the systems in question are 01, 07, and 11. The first is a small, wholly rural district with great wealth and very high MEL but very low MFI, the other two districts are industrial-urban with MFI in the high middle range, but low MEL. Teachers in the industrial-urban settings were relatively strong in reading, while teachers in the rural setting were not strong in reading. With only five systemic observations per cell, which way these three systems were grouped of course had a substantial influence on whether a significant interaction or a significant main effect or both appeared for reading.

The final analysis in the series was conducted as a cross-check on the urbanization dimension. Percent of the population in manufacturing was used to represent the urbanization dimension of communities and MEL the socio-economic dimension. Each of the variables in question were dichotomized, and the analysis run for intermediate teachers only. The results indicated that the main effects of median education level for both arithmetic and science were similar to those in the MEL by wealth analysis, although slightly weaker, but that there were no other significant effects.

Relationship of Estimated Socio-economic Status of Schools to Teacher Task and Scale Scores, with Teachers as Observations. If the socio-economic status of communities is a factor in teacher performance, as suggested above, the possibility must also be recognized that the socio-economic status of schools within communities might also be a factor in performance. Indeed, if a disproportionate numberof middle class or working class schools were identified in the sample with a particular type of community, differences between communities or districts might be shown to be a function of the class status of the schools involved rather than of the communities as a whole. To check this possibility, the distribution of schools of middle class, mixed, and working class student composition in districts classified MFI by wealth, and classified by median education level were tabled. The distribution may be examined in Tables 5-4 and 5-4a respectively.

As may be noted in the tables cited above, the distribution of schools by socio-economic status was relatively well balanced over the various cells, although slight disproportionality appeared in the MEL cells. To examine the effects on performance of the socio-economic status of the school in which the teacher taught, all teachers who had been in their current school for two years or less were excluded, as were special teachers and teachers in departmentalized situations. Teachers were subsequently

TABLE 5-4. Distribution of Schools by Socio-economic Status for Wealth by MFI Cells.*

Wealth	MFI					
		Hi			Lo	
	Middle	Mixed	Working	Middle	Mixed	Working
	03	03	03		01,01	01
	05	05	05			08
Hi			07,07,07	16	16,16	16,16
	10	10	10,10	15		
	11	11	11			
	13,13	13,13,13				
	13,13					
		04	04			02
Lo	09,09			06		06
	14			12	12	12
		17	17		18	
	19		19	20	20	20
Totals	12	9	11	5	7	8

TALBE 5-4a. Distribution of Schools by Socio-economic Status for Median Education Level Cells.*

	MEL					
		Hi			Lo	
	Middle	Mixed	Working	Middle	Mixed	Working
		01,01	01			07,07,07
	03	03	03			08
	05	05	05	11	11	11
	10	10	10,10	15		
	13,13			16	16,16	16,16
	13,13	13,13,13				
		04	04			02
	09,09			06		06
	14			12	12	12
		17	17		18	
	19		19	20	20	20
Totals	11	10	8	6	6	11

---

*The numbers in the cells are the I.D. numbers of the systems sampled.

separated according to grade range, primary and intermediate, and according to the median education level of the district. Task and scale scores were then examined within each MEL cell across the three levels of socio-economic status by analysis of variance.

The results of the analysis indicated that there were no significant differences on any task or scale among either primary or intermediate teachers by the socio-economic status of their school in the hi MEL cell. In the low MEL cell, Stability scores were significantly lower among primary teachers in mixed schools, but significantly higher among intermediate teachers in these same schools. In absence of other results, the extremely inconsistent results for Stability were beyond interpretation, and were considered by the writer to be too weak to warrant the introduction of school SES as a control variable. Hence, in subsequent analysis bearing of differences in score levels between system types, this variable was disregarded.

Quite aside from the largely statistical implications of the SES of schools discussed above, the fact that no differences attributable to socio-economic status was found, distinctly suggested that in systems of the size appearing in the sample, the kind of neighborhood in which the school was located made no difference in the quality of its teachers, as estimated by the instruments employed in the project.

Relationships of Median Family Income of School Districts and Length of Teacher Service in Districts to Teacher Task and Scale Scores, with Teachers as Observations. This series of analyses utilized only the scores of teachers in self-contained classrooms for whom complete data was available. The population differs from that employed in the preceding analyses in that it excludes teachers engaging in departmentalized instruction, but includes those with two or less years of experience in self-contained classrooms in their system. The restriction of the population to teachers in self-contained classrooms permitted the use of a combined or total Z score for arithmetic, reading and science as an estimate of over-all strength in task performance.

In the first analysis in the series, school districts were dichotomized on median family income, and primary and intermediate teachers separated. Two group analysis of variance was then utilized to examine differences in task performance and characteristics between high and low MFI districts according to grade range taught. The results for intermediate teachers, shown in Table 5-5 and for primary teachers, shown in Table 5-6, are quite similar to those in Table 5-3, although there are sharper differences between the high and low MFI groups in reading, and less sharp differences in Warmth-spontaneity among intermediate teachers in Table 5-5 than in Table 5-3. Thus, changes in the definition of the population appeared to have relatively little influence on the effects associated with median family income in the school districts.

TABLE 5-5. Differences in Task and Scores for Intermediate Teachers by Median Family Income in the School District.

Tasks and Scales	MFI	n	Mean	S.D.	F	P
Read.	Hi	136	103.50	15.71	17.44	.001
	Lo	97	94.80	15.61		
Arith.	Hi	136	103.50	13.58	13.84	.001
	Lo	97	96.45	13.87		
Sci.	Hi	136	52.10	9.58	9.52	.01
	Lo	97	48.11	9.52		
Z Total	Hi	136	258.90	25.77	31.21	.001
	Lo	97	239.37	27.02		
W-S	Hi	136	50.52	10.26	3.86	.10
	Lo	97	47.93	9.36		
Org.	Hi	136	49.61	9.75	0.04	n.s.
	Lo	97	49.32	10.29		
Vpt.	Hi	136	49.05	10.42	0.87	n.s.
	Lo	97	50.31	9.62		
Stab.	Hi	136	49.46	10.18	0.01	n.s.
	Lo	97	49.57	9.84		
Involve.	Hi	136	49.13	9.84	0.01	n.s.
	Lo	97	49.21	10.74		

TABLE 5-6. Differences in Task and Scale Scores for Primary Teachers by Median Family Income in the School District.

Tasks and Scales	MFI	n	Mean	S D	F	P
Read.	Hi	186	101.03	14.63	2.74	n.s.
	Lo	109	98.06	15.37		
Arith.	Hi	186	99.86	15.04	0.31	n.s.
	Lo	109	100.87	15.15		
Sci.	Hi	186	49.81	10.86	1.19	n.s.
	Lo	109	51.15	8.94		
Z Total	Hi	186	250.70	26.42	0.23	n.s.
	Lo	109	249.17	27.11		
W-S	Hi	186	50.19	8.77	0.28	n.s.
	Lo	109	49.60	10.03		
Org.	Hi	186	50.46	9.31	2.08	n.s.
	Lo	109	48.84	9.23		
Vpt.	Hi	186	49.49	10.29	0.00	n.s.
	Lo	109	49.44	8.98		
Stab.	Hi	186	50.23	9.07	0.68	n.s.
	Lo	109	49.84	10.04		
Involve.	Hi	186	49.86	9.13	.83	n.s.
	Lo	109	48.65	13.40		

For the second analysis in the series the dichotomy between high and low MFI districts and the separation of primary and intermediate teachers were maintained, then an additional separation of teachers divided according to whether they had been in their present system five or fewer years or six or more years was introduced. The point in separating teachers with 0-5 and six or more years of experience in their system was to gain estimates of the extent to which selective factors as opposed to behavior modification factors operated in producing differences between systems. If significant differences between high and low median income districts appeared for teachers who have five or fewer years of experience in these systems, evidence would be provided that selective factors may be operating, since these teachers are relatively new to their systems. On the other hand, if differences arose primarily between teachers who have been in their system six or more years, but there were no differences between systems among teachers 0-5 years of experience in their system, evidence would be provided that either behavior modification or differential selection following employment, or both, were operating. The possibility that there are differences between systems among teacher 0-5 years of experience in their system but not among teachers six or more years, and the possibility that there are differences between systems among both teachers of 0-5 and six or more years of experience in their system must also be recognized. In the event one or the other of the latter possibilities obtained, additional appraisals would be necessary to determine the relative effects of selection versus behavior modification factors.

The significance of the mean differences between teachers 0-5 and six or more years of experience in high versus low MFI districts may be observed in Tables 5-7 and 5-8 for intermediate and for primary teachers respectively. In reading these tables, differences between high MFI and low MFI districts for teachers in the same experience range (0-5, or six plus) may be examined by noting the means, F ratios and probability levels given in the columns. Differences between teachers in the same type of system but having different levels of experience in that type of system may be examined by examining the means,F ratios and probability levels in the rows.

Examination of Table 5-7 and 5-8 indicates that there are rather sharp differences among the various dependent variables by rows and columns. In reading, among intermediate teachers, it is apparent that there are highly significant differences between high and low MFI districts among teachers 0-5 years of experience, but less difference among those six plus years, with a collateral tendency for teachers six plus years to perform significantly less strongly than their less experienced colleagues in high MFI districts. Among primary teachers these effects are largely absent, although their is a slight tendency for performance to drop as experience in either high or low MFI districts increases. In arithmetic, differences between intermediate teachers are attributable to differences in the MFI of the districts, but not to experience, while among primary

TABLE 5-7. Differences in Task and Scale Scores for Intermediate Teachers by Years of Experience in their System and Median Family Income in the School District.

Tasks and Scales	MFI F and P	0-5 Years Mean	S.D.	6 or more Yrs. Mean	S.D.	F	P
	Hi	106.40	13.43	100.21	17.61	5.35	.05
Read.	Lo	96.71	15.20	93.15	15.91	1.26	n.s.
	F(p)	13.02	(.01)	4.98	(.05)		
	Hi	104.90	12.82	101.71	14.03	1.90	n.s.
Arith.	Lo	98.02	14.65	95.10	13.16	1.07	n.s.
	F(p)	7.13	(.025)	6.71	(.05)		
	Hi	52.25	9.81	51.94	9.46	0.03	n.s.
Sci.	Lo	49.27	8.49	47.12	10.92	1.15	n.s.
	F(p)	2.83	(n.s.)	6.43	(.025)		
	Hi	263.69	22.64	253.86	28.18	5.04	.05
Z Total	Lo	244.00	26.09	235.37	27.41	2.50	n.s.
	F(p)	18.62	(.01)	12.57	(.01)		
	Hi	52.13	7.52	48.25	12.35	5.48	.05
W-S	Lo	49.18	10.39	46.87	8.32	1.14	n.s.
	F(p)	3.58	(n.s.)	0.48	(n.s.)		
	Hi	48.79	8.28	50.30	11.12	0.81	n.s.
Org.	Lo	48.58	10.10	49.96	10.51	0.43	n.s.
	F(p)	0.01	(n.s.)	0.03	(n.s.)		
	Hi	48.39	9.64	50.11	11.09	0.93	n.s.
Vpt.	Lo	49.20	8.74	51.27	10.30	1.12	n.s.
	F(p)	0.21	(n.s.)	1.29	(n.s.)		
	Hi	50.46	9.11	48.05	11.08	1.92	n.s.
Stab.	Lo	48.87	11.27	50.17	8.47	0.42	n.s.
	F(p)	0.07	(n.s.)	1.29	(n.s.)		
	Hi	47.24	9.60	51.08	10.09	5.13	.05
Involve.	Lo	49.27	11.83	49.15	9.81	0.03	n.s.
	F(p)	1.03	(n.s.)	1.06	(n.s.)		

TABLE 5-8. Differences in Task and Scale Scores for Primary Teachers by Years of Experience in their System and Median Family Income in the School District.

Tasks and Scales	MFI F and P	0-5 Years Mean	0-5 Years S.D.	6 or more Yrs. Mean	6 or more Yrs. S.D.	F	P
	Hi	103.06	13.41	98.96	15.57	3.71	n.s.
Read.	Lo	100.79	15.57	95.46	15.28	3.34	n.s.
	F(p)	0.88	(n.s.)	1.77	(n.s.)		
	Hi	99.27	14.72	100.46	15.42	0.29	n.s.
Arith.	Lo	99.94	15.49	101.75	14.19	0.39	n.s.
	F(p)	0.07	(n.s.)	0.25	(n.s.)		
	Hi	51.51	10.76	48.08	10.48	4.74	.05
Sci.	Lo	51.34	7.92	50.98	9.89	0.05	n.s.
	F(p)	0.01	(n.s.)	0.05	(n.s.)		
	Hi	253.84	24.46	247.49	27.60	2.71	n.s.
Z Total	Lo	252.08	23.03	246.41	30.43	1.19	n.s.
	F(p)	0.18	(n.s.)	0.05	(n.s.)		
	Hi	50.57	9.25	49.79	8.28	0.37	n.s.
W-S	Lo	53.62	9.51	45.79	9.03	19.48	.001
	F(p)	3.60	(.10)	7.62	(.01)		
	Hi	51.37	8.98	49.52	9.59	1.85	n.s.
Org.	Lo	50.30	8.93	47.46	9.38	2.61	n.s.
	F(p)	1.84	(n.s.)	1.07	(n.s.)		
	Hi	48.79	6.91	50.21	11.01	0.08	n.s.
Vpt.	Lo	46.77	7.74	52.34	9.25	10.47	.01
	F(p)	1.84	(n.s.)	2.74	(n.s.)		
	Hi	50.49	9.71	49.96	8.41	0.15	n.s.
Stab.	Lo	51.30	9.36	47.38	10.37	4.29	.05
	F(p)	0.24	(n.s.)	2.74	(n.s.)		
	Hi	47.80	9.71	51.78	8.52	8.05	.01
Involve.	Lo	48.21	12.71	49.07	14.12	0.11	n.s.
	F(p)	0.02	(n.s.)	2.13	(n.s.)		

TABLE 5-8a. Number of Teachers in each Group in Tables 5-7 and 5-8.

		0-5 Years	6+ Years
Hi MFI	Intermediate	N=72	N=63
	Primary	N=94	N=92
Lo MFI	Intermediate	N=45	N=52
	Primary	N=53	N=56
Totals		N=264	N=263

teachers, there are no significant effects for either system type or experience. In science, the more experienced teachers in high MFI districts are slightly superior to those in low MFI districts, a result which appears to arise from a slight downward shift in the performance levels of the more experienced teachers in low MFI districts compared to the less experienced teachers in these districts. Among primary teachers, this downward shift occurs significantly among those in high MFI districts, a result not altogether congruous with that for intermediate teachers.

The differences among the various intermediate groupsfor Z total reflect, of course, the component differences, i.e. reading, arithmetic and science; however, they also bring into sharper focus total task performance differences between high and low median income districts, especially in the group 0-5 years of experience in their district. These differences clearly suggest that there is a separation between high and low districts in the task performance strength of intermediate teachers recently employed. These differences appear to be maintained among the more experienced group, but not quite to the same degree. Some slippage appears to occur. Among primary teachers, on the other hand, there is no evidence of differences in task performance among teachers recently employed in the systems, and little evidence of significant slippage. Thus it appears to be the case that the selection criteria which operate for intermediate teachers do not operate, or do not operate as fully, for primary teachers.

With respect to the personal-social scales from the Characteristics Schedule, the situation is quite different. Among intermediate teachers, there is evidence of slippage between the 0-5 and six plus years group in high MFI districts on Warmth-spontaneity, while at the same time there is a slight upward shift in the same group on Involvement, but there is no additional evidence of differences by either experience or type of district on the personal-social variables among intermediate teachers. Among primary teachers, on the other hand, there is a very great evidence of slippage between teachers of 0-5 and six plus years in their system in low MFI districts in Warmth-spontaneity, Viewpoint, and to a lesser degree, Stability, while the Involvement scores of primary teachers in high MFI districts show a substantial upward shift. These differences, however, are associated with differences between systems only in the six plus years group on Warmth-spontaneity. This particular arrangement of differences tentatively suggests that personal-social characteristics are approximately equally distributed across systems among primary teachers relatively recent to their system, but that low MFI districts are either unable to retain teachers beyond the fifth year of service high in Warmth-spontaneity, low in Viewpoint (child-centered) and high in Stability, or else that these systems modify the behavior of their teachers on these factors in a direction which, from the viewpoint of the author at least, is the reverse of what one would expect, i.e. toward less Warmth-spontaneity, greater subject-centeredness and less emotional Stability.

An important adjunct to the interpretation of the patterns of differences described above may be observed in Table 5-8a in which the number of teachers associated with each cell in Tables 5-7 and 5-8 are given. The most significant feature of Table 5-8a is that almost exactly 50 percent of the teachers in the study had been in their school system five years or less. These data not only signify that the apparent turnover rate for elementary teachers in these systems is quite high, but an addition make abundantly clear the possibility that differences between teachers 0-5 and six plus years in their system can come about by differential attrition, which encompasses both selection procedures within systems subsequent to employment, or "weeding out," and teacher self-selection, or differential teacher mobility.

Relationship of Systemic Norms to Task and Scale Scores. In both the teacher behavior modification model and the differential attraction model discussed in Chapter 3, norms for teacher behaviors or characteristics were viewed as important components, although in each model these norms were viewed as operating in different ways. Two measures of systemic norms were used in the study, the Teacher Behavior Weighting Scales (TBWS) and teacher ratings.

As a preliminary analysis, the TBWS scores of the principals who completed this scale were examined in relation to five variables: the size of the school (enrollment), the experience of the principal, the grade level at which the principal had taught before becoming a principal, the hours of work he had completed beyond the Master's Degree, and the socio-economic status of the school. The analyses for the first four of the variables above revealed only one significant difference: principals with more than fifteen years of experience were higher on TBWS Scale A (managment, control), (F=5.47,2and 40 df., p .01) than were the less experienced principals. Analysis by the socio-economic status of the school in which the principals was located revealed, however, quite consistent differences, as may be viewed in Table 5-9. In the latter table one may note that both the A and B scales follow an inverted U-shaped function over socio-economic status, while the C and D scales follow a U-shaped function. Since the A and B scales are non-ipsative with respect to the C and D scales, a reversal in the functions would of course be expected.

The fact that principals in schools of mixed socio-economic composition scored significantly higher on both the A and B scales of the TBWS distinctly implied that if these scores truly reflected the norms of the principals, and if the norms of principals at all influenced the level of teacher performance, teachers in schools with mixed socio-economic status should show higher performance either on the Total Z score for the teaching tasks, or on one or more of the component scores, i.e. reading, arithmetic or science. As earlier observed, however, there were no differences in task scores between schools classified by SE status, when the median education level of the community was controlled. The validity of the TBWS as guaged against this criterion therefore fell into doubt.

Two more direct tests of the validity of the TBWS were subsequently employed. In the first, a chi square analysis of the distribution of principals whose A and B scores were higher than their C plus C scores was tested against high versus low MFI districts. Since the teaching task scores of intermediate teachers in high MFI districts were known to be significantly different from those in low MFI districts, a disproportionate number of principals with high A plus B scores should have appeared in the high MFI districts. There was, however, no significant disproportionality present (X^2=.374, 1 df).

For the second analysis, teachers were separated by grade range, primary and intermediate, and by MFI, high and low. Within these groups, they were then divided according to whether the principal of their school was task-oriented (high A plus B) or social-emotionally oriented (high C plus D), and the task and scale scores of teachers in the resulting groups examined by analysis of variance. Only one significant difference appeared in the analyses; primary teachers in high MFI districts who had task oriented principals showed significantly higher Involvement scores (F=6.34, p .05) than teachers in the same systems with social-emotionally oriented principals. With respect to the validity of the TBWS, this result was regarded predominantly as an irrelevancy, and the inference was made that the TBWS scores of principals were unrelated to the performances of their teachers.

The second series of analyses relating systemic norms to teacher task and scale performance was conducted using principal's ratings of their teachers as the definition of systemic norms. Under this definition, differences in performance or characteristics between teachers ranked in

TABLE 5-9. Mean Scores of Principals on Four TBWS Scales by the Socio-economic Status of their Schools.

		S-E Status				
Scale		Middle class N=13	Mixed N=14	Working class N=17	F	P
A Management, control	Mean	9.46	17.07	7.11	7.63	.01
	SD	4.63	10.47	5.43		
B Learning	Mean	10.08	13.78	8.06	4.07	.025
	SD	5.99	6.73	4.07		
C Emotional	Mean	9.92	6.79	11.35	3.47	.05
	SD	6.42	3.44	4.44		
D Social	Mean	8.00	6.00	10.35	3.86	.05
	SD	4.93	4.19	4.03		

the upper two versus the lower two of four forced-choice categories are taken to signify the presence of a valued or preferred task performance or characteristic. If there are no differences on a particular characteristic between teachers ranked in the categories in question, the characteristic is taken to be non-valued, non-preferred, or as not representing a norm.

The analyses in which principal's rating were used as the definition of norms were conducted specifically with the behavior modification model in mind. Under this model, if a norm can be shown to be present in a particular type of school district, one condition of the model is met. If in addition, there is evidence of a shift in performance among teachers as experience in the system increases, and this shift is in the direction of the norm, evidence is provided that the presence of the norm may be an influence on the behavior of the teachers under it. On the basis of analyses shown on preceding pages, and especially in Table 5-7 and 5-8, however, there is no reason to suppose that the model thus conceptualized could be found completely workable. Quite simply, the difficulty lies in the fact that on only one characteristic, Involvement, is there any evidence of an increase in performance. On the other hand, there are instances of no significant shift in performance with experience, as well as instances of a significant downward shift in performance.

The fact that downward shift occured among certain teacher performances and characteristics in some districts and not in others introduced the possibility that the effect of norms might not be to increase performance or to change characteristics, but rather to maintain them at a particular level against countervailing variables, such as differential mobility, which would, without the presence of the norm, create a downward shift. Under this interpretation, if the performances or characteristics of teachers either remained the same or increased with experience, and a norm were shown to be present, evidence for the influence of the norm on the behavior would be provided. Collaterally, if no norm could be shown to be present and there was a significant downward shift in performance, the countervailing variables would be interpreted as having an uninterupted effect. If, however, there is no norm present and there is also no change in the performances or characteristics with experience, there is evidence that systemic norms are not relevant to performances or characteristics of teachers.

Holding the above possibilities in mind, primary and intermediate teachers were separated, and within each group teachers 0-5 and six plus years of experience in their systems were divided. These groups were then divided into four groups, each group representing one system type derived by dichotomizing on per pupil wealth and MFI. With each of these groups, two separate analyses were conducted. In the first, teachers 0-5 versus six plus years of experience were compared on each task and scale. In the second, teachers rated high (rank 1 and 2) versus low (ranks 3 and 4) were compared. The first analysis is for score shift, the second for the presence of norms. It should be noted that the analysis for score shift is based on the same number of teachers as the analysis in Tables 5-7 and 5-8, and differs from this analysis only in that within each MFI category (high, low) systems are divided by wealth. The analysis for norms, however, is based on fewer teachers than the analysis for score shift, since, as noted in Chapter 4, three principals failed to rate their teachers.

The results of the analysis are shown in Tables 5-10 through 5-13. In reading these tables, the differences between teachers 0-5 and six plus years of service to their system may be found in the first column, and the results of the rating in the second column. In these tables there is no uniform correspondence between whether or not a particular performance or characteristic is significantly valued, as expressed by differences between teachers rated high versus low, and shifts in task or scale scores. For example, in Table 5-10, intermediate teachers with higher scores in task performance in teaching reading are rated significantly higher than those with lower scores, collaterally, there is no significant decline in reading scores in this group, raising the possibility that the maintanence of reading performance score levels is associated with the presence of norms. In Table 5-13, on the other hand, there is also no significant downward shift in reading performance, and there is also no significant difference on the reading scores for teachers rated high versus low. Overall, there is no consistent evidence in these tables that the presence of norms in a particular type of system is associated with shifts in score levels or the maintanence of score levels as experience increases in systems of that type.

An aspect of Tables 5-10 through 5-13 which should not be overlooked is that differences between teachers rated high versus those rated low occur almost wholly within hi MFI - hi wealth districts. Thus, to the extent that norms are significantly present, they are identified largely with systems in which there is a conjunction of above average per pupil wealth and above average median family income.

Relationship of Teacher Task and Scale Scores to Type of Entry into Their System. The analyses on preceding pages strongly suggest that the behavior modification model proposed in Chapter 3 is not a workable model as examined within the limits of the present data. The alternative model is the differential attraction model. To explore the potential workability of the latter model, the kind of entry the teacher made into the system was examined. Three types of entry into the systems were recognized. First, a new entry, meaning a teacher taken directly from a college teacher preparation program and placed in the system within the past two years. Second, the transfer entry, meaning a teacher who had initially been employed elsewhere, but had at some time in the past, and most typically within the last ten years, transferred to her current system. Third, the teacher who had entered her current system as a new or beginning teacher, and had stayed on in the system. The latter group was labeled the "always in" group.

The scores of these three groups of teachers may be observed in Tables 5-14 and 5-15. In these tables, the scores for transfer and "always in" teachers are shown by district type (1, 2, 3, 4) grouped under Hi versus Lo MFI, while the scores for new teachers of whom there were relatively few in Type 3 and Type 4 districts, were pooled in Type 1 and 2, and Type 3 and 4 districts, and are shown centered under the Hi MFI and Lo MFI columns respectively. Single classification analysis of variance was used to test the differences among the means shown in each row in the tables. The number of teachers involved is shown in Tables 5-14a and 5-15a for intermediate and for primary teachers respectively.

In interpreting Tables 5-14 and 5-15 it is important to note, first that the ratio of new and transfer teachers to teachers always in their system is in excess of 3:1 as shown below Tables 5-14a and 5-15a. To put the matter another way, the chances of encountering a teacher who has always been in the system in which she started to teach is, in the present sample, always less than of one in four. In some instances, as among intermediate teachers in high MFI districts, the chances are less than one in five. Thus, the great preponderance of teachers in the systems sampled have either recently moved to the system or earlier transferred to it from some other system. Examining the means now for intermediate teachers in the hi MFI districts (Table 5-14), it is immediately apparent that the wide

TABLE 5-10. Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6+ Years in their System and Relationships of Scores to Ratings (norms) in Hi MFI - Hi Wealth (Type 1) Districts.

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Intermediate								
Read.	0-5	104.98			Hi	106.85		
	6+	101.65	1.07	n.s.	Lo	97.09	8.38	.01
Arith.	0-5	104.40			Hi	103.91		
	6+	102.35	0.51	n.s.	Lo	100.68	1.21	n.s.
Sci.	0-5	51.69			Hi	51.87		
	6+	51.42	0.02	n.s.	Lo	51.31	0.06	n.s.
Z Total	0-5	261.06			Hi	262.63		
	6+	255.42	1.08	n.s.	Lo	249.39	5.72	.05
W-S	0-5	52.56			Hi	51.08		
	6+	48.65	3.46	n.s.	Lo	50.36	0.11	n.s.
Org.	0-5	47.90			Hi	49.22		
	6+	50.05	0.95	n.s.	Lo	48.12	0.23	n.s.
Vpt.	0-5	48.40			Hi	49.80		
	6+	48.72	0.02	n.s.	Lo	47.95	0.01	n.s.
Stab.	0-5	50.25			Hi	50.24		
	6+	48.42	0.64	n.s.	Lo	48.73	0.38	n.s.
Involve.	0-5	46.31			Hi	50.54		
	6+	50.67	4.16	.05	Lo	46.54	3.13	n.s.

TABLE 5-10. (continued)

Tasks and Scales	Years in System	Mean	F	P	Rating (norms	Mean	F	P
Primary								
Read.	0-5	102.82			Hi	101.98		
	6+	101.60	0.20	n.s.	Lo	102.46	0.02	n.s.
Arith.	0-5	97.60			Hi	100.98		
	6+	102.17	3.31	n.s.	Lo	99.78	0.05	n.s.
Sci.	0-5	51.13			Hi	50.30		
	6+	48.38	2.19	n.s.	Lo	50.87	0.05	n.s.
Z Total	0-5	251.53			Hi	252.67		
	6+	252.16	0.02	n.s.	Lo	253.11	0.01	n.s.
W-S	0-5	50.24			Hi	51.71		
	6+	49.93	0.03	n.s.	Lo	47.28	6.17	.025
Org.	0-5	51.11			Hi	51.63		
	6+	49.35	1.06	n.s.	Lo	47.33	5.66	.05
Vpt.	0-5	49.09			Hi	47.35		
	6+	49.67	0.13	n.s.	Lo	52.13	6.07	.025
Stab.	0-5	50.18			Hi	51.25		
	6+	49.91	0.13	n.s.	Lo	48.02	3.75	n.s.(.10)
Involve.	0-5	48.49			Hi	51.25		
	6+	51.10	2.19	n.s.	Lo	45.76	11.61	.01

TABLE 5-11. Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6+ Years in their System and Relationships of Scores to Rating (norms) in Hi MFI - Lo Wealth (Type 2) Districts.

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Intermediate								
Read.	0-5	109.25			Hi	107.44		
	6+	97.10	6.44	.025	Lo	103.32	0.79	n.s.
Arith.	0-5	105.92			Hi	104.06		
	6+	100.35	1.94	n.s.	Lo	103.11	0.06	n.s.
Sci.	0-5	53.38			Hi	53.72		
	6+	53.05	0.01	n.s.	Lo	50.79	1.18	n.s.
Z Total	0-5	268.96			Hi	265.22		
	6+	250.50	6.63	.025	Lo	257.74	1.43	n.s.
W-S	0-5	51.83			Hi	52.28		
	6+	47.40	1.99	n.s.	Lo	46.84	2.48	n.s.
Org.	0-5	50.58			Hi	47.33		
	6+	50.85	0.02	n.s.	Lo	52.78	4.63	.05
Vpt.	0-5	48.38			Hi	47.89		
	6+	53.10	2.14	n.s.	Lo	52.37	1.89	n.s.
Stab.	0-5	50.88			Hi	48.44		
	6+	47.25	2.08	n.s.	Lo	48.90	0.03	n.s.
Involve.	0-5	49.08			Hi	51.11		
	6+	51.95	1.02	n.s.	Lo	48.16	1.02	n.s.

TABLE 5-11. (continued)

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Primary								
Read.	0-5	103.53			Hi	98.30		
	6+	94.44	7.42	.01	Lo	99.91	0.18	n.s.
Arith.	0-5	102.50			Hi	101.59		
	6+	97.53	1.41	n.s.	Lo	99.09	0.29	n.s.
Sci.	0-5	52.81			Hi	48.22		
	6+	47.56	2.55	n.s.	Lo	51.14	0.73	n.s.
Z Total	0-5	258.31			Hi	248.11		
	6+	239.53	8.30	.01	Lo	250.14	0.07	n.s.
W-S	0-5	51.22			Hi	50.97		
	6+	49.56	0.77	n.s.	Lo	49.96	0.23	n.s.
Org.	0-5	51.88			Hi	50.92		
	6+	49.82	0.82	n.s.	Lo	50.36	0.05	n.s.
Vpt.	0-5	48.19			Hi	49.32		
	6+	50.97	1.05	n.s.	Lo	48.96	0.01	n.s.
Stab.	0-5	51.09			Hi	50.54		
	6+	50.03	1.05	n.s.	Lo	50.45	0.00	n.s.
Involve.	0-5	47.03			Hi	50.84		
	6+	52.94	9.04	.01	Lo	48.68	0.85	n.s.

TABLE 5-12. Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6+ Years in their System and Relationships of Scores to Ratings (norms) in Lo MFI - Hi Wealth (Type 3) Districts.

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Intermediate								
Read.	0-5	98.70			Hi	97.18		
	6+	96.21	0.27	n.s.	Lo	92.33	0.68	n.s.
Arith.	0-5	96.87			Hi	97.68		
	6+	93.46	0.78	n.s.	Lo	89.75	3.10	n.s.
Sci.	0-5	48.87			Hi	48.86		
	6+	46.00	1.21	n.s.	Lo	45.42	1.13	n.s.
Z Total	0-5	244.43			Hi	243.73		
	6+	235.67	1.22	n.s.	Lo	227.50	2.53	n.s.
W-S	0-5	52.83			Hi	49.23		
	6+	46.88	4.90	.05	Lo	50.00	0.05	n.s.
Org.	0-5	52.74			Hi	51.50		
	6+	50.13	0.86	n.s.	Lo	51.50	0.00	n.s.
Vpt.	0-5	49.39			Hi	51.09		
	6+	53.08	1.78	n.s.	Lo	52.20	0.25	n.s.
Stab.	0-5	52.13			Hi	52.46		
	6+	51.46	0.04	n.s.	Lo	49.04	0.74	n.s.
Involve.	0-5	51.39			Hi	48.14		
	6+	48.45	0.73	n.s.	Lo	51.00	0.51	n.s.

TABLE 5-12. (continued)

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Primary								
Read.	0-5	100.19			Hi	100.13		
	6+	98.30	0.18	n.s.	Lo	96.59	0.47	n.s.
Arith.	0-5	102.38			Hi	103.54		
	6+	100.04	0.35	n.s.	Lo	98.47	1.44	n.s.
Sci.	0-5	49.96			Hi	48.58		
	6+	49.03	0.00	n.s.	Lo	50.71	0.49	n.s.
Z Total	0-5	252.54			Hi	252.25		
	6+	247.65	0.43	n.s.	Lo	245.76	0.58	n.s.
W-S	0-5	53.08			Hi	48.90		
	6+	44.13	11.26	.01	Lo	47.18	0.27	n.s.
Org.	0-5	51.42			Hi	47.08		
	6+	45.70	3.96	.05	Lo	50.12	1.10	n.s.
Vpt.	0-5	48.77			Hi	49.60		
	6+	54.04	5.52	.05	Lo	53.59	2.82	n.s.
Stab.	0-5	51.15			Hi	48.58		
	6+	45.70	3.96	n.s.	Lo	49.41	0.07	n.s.
Involve.	0-5	46.89			Hi	48.75		
	6+	46.22	0.02	n.s.	Lo	43.59	0.09	n.s.

TABLE 5-13. Differences in Task and Scale Scores for Primary and for Intermediate Teachers 0-5 vs 6+ Years in their System and Relationships of Scores to Ratings (norms) in Lo MFI - Lo Wealth (Type 4) Districts.

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Intermediate								
Read.	0-5	94.64			Hi	92.88		
	6+	90.54	0.94	n.s.	Lo	91.39	0.11	n.s.
Arith.	0-5	99.23			Hi	99.23		
	6+	96.50	0.43	n.s.	Lo	96.50	0.43	n.s.
Sci.	0-5	49.68			Hi	48.28		
	6+	48.07	0.27	n.s.	Lo	49.67	0.19	n.s.
Z Total	0-5	243.55			Hi	236.94		
	6+	235.11	1.20	n.s.	Lo	242.17	0.43	n.s.
W-S	0-5	45.36			Hi	46.37		
	6+	46.86	0.34	n.s.	Lo	45.89	0.03	n.s.
Org.	0-5	44.27			Hi	46.75		
	6+	49.82	3.61	n.s.	Lo	48.44	0.29	n.s.
Vpt.	0-5	49.00			Hi	50.78		
	6+	49.71	0.06	n.s.	Lo	46.94	1.83	n.s.
Stab.	0-5	45.46			Hi	46.81		
	6+	49.07	2.41	n.s.	Lo	48.67	0.57	n.s.
Involve.	0-5	47.05			Hi	48.56		
	6+	49.75	0.93	n.s.	Lo	48.56	0.00	n.s.

TABLE 5-13. (continued)

Tasks and Scales	Years in System	Mean	F	P	Rating (norms)	Mean	F	P
Primary								
Read.	0-5	101.37			Hi	99.03		
	6+	93.48	4.41	.05	Lo	94.90	1.09	n.s.
Arith.	0-5	97.59			Hi	101.23		
	6+	102.94	1.60	n.s.	Lo	99.79	0.11	n.s.
Sci.	0-5	52.67			Hi	53.00		
	6+	52.15	0.05	n.s.	Lo	51.72	0.32	n.s.
Z Total	0-5	251.63			Hi	250.03		
	6+	245.55	0.69	n.s.	Lo	246.41	0.24	n.s.
W-S	0-5	54.15			Hi	50.52		
	6+	46.94	8.98	.01	Lo	49.83	0.07	n.s.
Org.	0-5	49.22			Hi	48.68		
	6+	48.70	0.04	n.s.	Lo	49.35	0.10	n.s.
Vpt.	0-5	44.70			Hi	48.68		
	6+	50.64	6.51	.05	Lo	47.21	0.37	n.s.
Stab.	0-5	51.44			Hi	49.17		
	6+	48.54	1.20	n.s.	Lo	50.00	0.01	n.s.
Involve.	0-5	49.48			Hi	52.13		
	6+	51.06	0.35	n.s.	Lo	48.45	1.95	n.s.

TABLE 5-14. Mean Scores of Intermediate Teachers New to their Current System, who Transferred to their System and who have Always been in their System, by MFI and by District Type

Tasks and Scales		Hi MFI 1	2	Lo MFI 3	4	F	P
Read.	New	105.18		100.82			n.s.
	Trans.	102.34	105.89	95.12	88.45	7.07	.01
	Always in	103.38	99.88	97.67	100.25	0.25	n.s.
Arith.	New	106.03		101.00			n.s.
	Trans.	103.82	101.07	95.12	95.42	3.46	.05
	Always in	100.31	98.25	96.22	100.08	0.17	n.s.
Sci.	New	53.42		49.82			n.s.
	Trans.	51.04	51.18	48.92	46.13	1.68	n.s.
	Always in	49.56	56.12	50.44	51.25	0.72	n.s.
W-S	New	53.42		50.00			n.s.
	Trans.	48.09	48.11	47.04	47.29	0.08	n.s.
	Always in	53.06	55.11	48.11	49.00	1.54	n.s.
Org.	New	48.02		49.41			n.s.
	Trans.	49.79	49.03	50.48	47.97	0.27	n.s.
	Always in	50.38	50.38	52.79	53.41	0.54	n.s.
Vpt.	New	46.57		46.88			n.s.
	Trans.	49.25	51.11	52.04	48.71	0.71	n.s
	Always in	50.38	50.38	52.79	53.41	0.22	n.s.
Stab.	New	50.53		46.06			n.s.
	Trans.	48.54	47.89	51.04	50.61	0.71	n.s.
	Always in	53.50	48.50	50.22	46.75	1.22	n.s.
Involve.	New	46.02		48.64			n.s.
	Trans.	51.14	49.63	49.40	50.35	0.20	n.s.
	Always in	49.25	46.62	49.33	47.50	0.15	n.s.

Table 5-14a. Number of Subjects in each Cell in Table 5-14

	Hi MFI 1	2	Lo MFI 3	4
New	N=38		N=17	
Trans.	N=44	N=27	N=25	N=31
Always in	N=16	N=8	N=9	N=12

(Ratio of new and transfer to always in, Hi MFI = 109/24; Lo MFI = 73/21)

TABLE 5-15. Mean Scores of Primary Teachers who are New to their Systems, who Transferred to their System and who have Always been in System, by MFI and District Type.

Tasks and Scales		Hi MFI		Lo MFI			
		1	2	3	4	F	P
Read.	New	105.04		102.95			n.s.
	Trans.	99.46	95.08	100.38	94.95	1.21	n.s.
	Always in	106.94	101.10	96.42	97.54	2.32	n.s.
Arith.	New	99.08		99.20			n.s.
	Trans.	97.22	96.40	101.15	100.85	1.08	n.s.
	Always in	106.81	106.40	98.75	104.85	0.89	n.s.
Sci.	New	51.16		52.50			n.s.
	Trans.	48.88	49.05	49.54	50.83	0.31	n.s.
	Always in	51.07	50.80	49.67	51.92	0.12	n.s.
W-S	New	50.92		55.70			n.s.
	Trans.	49.40	50.10	45.07	47.29	1.75	n.s.
	Always in	50.93	46.70	51.92	48.54	0.84	n.s.
Org.	New	49.32		51.00			n.s.
	Trans.	49.50	51.48	47.46	47.51	1.34	n.s.
	Always in	52.58	47.60	47.58	49.14	1.88	n.s.
Vpt.	New	49.68		46.50			n.s.
	Trans.	50.72	50.18	54.08	48.88	1.28	n.s.
	Always in	46.77	51.70	47.67	47.46	0.68	n.s.
Stab.	New	50.04		53.55			n.s.
	Trans.	49.38	51.23	47.12	47.34	1.42	n.s.
	Always in	50.84	47.60	46.08	51.92	0.92	n.s.
Involve.	New	48.12		48.20			n.s.
	Trans.	49.32	52.75	47.92	49.00	1.20	n.s.
	Always in	51.29	46.10	42.75	52.39	3.14	n.s.

TABLE 15a. Number of Subjects in each Cell in Table 5-15.

	Hi MFI		Lo MFI	
	1	2	3	4
New	N=25		N=17	
Trans.	N=68	N=40	N=26	N=41
Always in	N=31	N=10	N=12	N=13

(Ratio of new and transfer to the system to always in, Hi MFI = 133/41, Lo MFI = 87/25) performance advantage shown by these systems in reading, arithmetic, and science in Tables 5-3 and 5-5, arise almost wholly from relatively small advantages gained from the new teacher group and the transfer teacher group, with perhaps a slight advantage arising from the "always in" group, although the small number of teachers in each type of district in the latter group make clear that the advantage gained from this source cannot be very great.

A similar state of affairs exists for Warmth-spontaneity, but in this instance, the intermediate teachers transferring to high MFI districts are relatively weak, and the new teachers entering the system together with the teachers who have always been in the system, do not contribute sufficient weight to the over-all mean for strong effects to show. As may be observed in Table 5-5, the effects for W-S are significant only at p=.10.

In appraising the means for primary teachers, as shown in Table 5-15, it is important to bear in mind that in preceding tables, no over-all differences between high and low MFI appeared for these teachers. There were however, differences between high and low MFI teachers with six plus years of service in their system, with these differences apparently attendent upon the slippage of scores in low MFI districts in Warmth-spontaneity, Viewpoint, and Stability, as shown in Table 5-8, and Tables 5-12 and 5-13. The source of slippage on the personal-social scales in these systems appears to lie in part in the unusually high score means of their new teachers, in contrast to the typically lower mean scores of teachers who have transfered to the system or have always been in it. The level of scores for new teachers in these systems, together with the fact that teachers 0-5 years of service in these systems are markedly more Warm-spontaneous and child-centered than their colleagues with greater service, strongly suggests that both differential attraction and differential attrition operate in these systems. The systems attract high quality new teachers, but less high quality experienced teachers . Subsequently, from these groups they appear to lose the more Warm-spontaneous and more child-centered teachers, producing a distinct downward shift in the scores as experience increases.

Relationships between Systemic Norms and System Types. The analyses shown on preceding pages have strongly suggested that there is no relationship between systemic norms and the level of teacher performance or characteristics in the system. These analyses did not provide full information, however, on the relationships between the principal factors associated with communities, i.e. per pupil wealth and MFI, and the TBWS and ratings as norms.

A factor to be taken into account in the above relationships is that principals had a decided tendency to place beginning or new teachers in the lowest rank in the ratings, probably as a consequent of the forced choice rating which required that someone be placed in the lowest rank.

On the average, however, new teachers show higher performance and higher score levels on the personal-social scales than do more experienced teachers, and when placed in the lowest rank on the ratings have the effect of greatly accelerating the variances within that group. To correct for this factor, teachers with two or less years of experience were deleted from the analyses described in this section.

For the first analysis, primary and intermediate teachers were separated by high and low MFI districts, then divided according to whether the principal of the school appeared as task oriented or social-emotionally oriented on the TBWS. The scores of teachers rated high versus those rated low were then analyzed within each cell. The results may be observed for high MFI districts in Table 5-16 and for low MFI districts in Table 5-17.

The results indicate that in high MFI districts, high performance in teaching reading is preferred among intermediate teachers, while high involvement is preferred among primary teachers, irrespective of the principal's TBWS scores. There is some tendency, however, for a differential effect to occur on Organization, with principals high in social-emotional orientation preferring intermediate teachers low in this characteristic, while high task oriented principals place no particular value on the characteristic . Among both primary and intermediate teachers under high task oriented principals there is also a weak but consistent tendency for principals to prefer or rate higher those teachers with greater Warmth-spontaneity.

In low MFI districts, principals who were high in social-emotional orientation preferred the more highly involved primary teachers, but there were no other significant effects.

For the second analysis, primary and intermediate teachers were divided according to per pupil wealth of their district, but the TBWS scores of principals were ignored. Teachers rated high versus those rated low were then compared within each of the resulting four cells. The results may be observed in Table 5-18.

The results for this analysis have a very distinct structure. In high wealth districts, value appears to be attached to performance in reading and arithmetic, and to a much less degree in science, among intermediate teachers, resulting in distinct differences in Z Total for teachers rated high versus low. To put the matter another way, using ratings as the definition of systemic norms, principals in high wealth districts show a distinct task orientation with respect to intermediate teachers. There is, however, no attribution of significance to the personal-social variables among intermediate teachers. The reverse pattern appears for primary teachers, for whom task performance appears to be insignificant, but for whom Warmth-spontaneity, child-centeredness (low Viewpoint score) Stability and Involvement all appear as criterial or preferred attributes.

TABLE 5-16. Relationships between TBWS Scores as Norms and Ratings as Norms, Intermediate and Primary Teachers, Hi MFI Districts.

Tasks and Scales	High-Task Oriented, Rating				High Social-emotional Oriented, Rating			
	Hi	Lo	F	P	Hi	Lo	F	P
Intermediate	N=28	N=33			N=24	N=23		
Read.	108.39	99.42	5.47	.05	107.38	95.30	9.11	.01
Arith.	107.96	104.09	1.58	n.s.	102.08	96.13	2.29	n.s.
Sci.	55.46	52.15	1.84	n.s.	50.67	50.52	0.00	n.s.
Z Total	271.82	255.67	7.32	.01	260.13	241.96	7.51	.01
W-S	54.18	49.51	3.16	.10	47.92	49.61	0.47	n.s.
Org.	50.93	48.54	0.87	n.s.	46.04	51.34	4.89	.05
Vpt.	51.11	48.36	1.00	n.s.	47.58	50.70	1.40	n.s.
Stab.	52.82	48.24	3.22	.10	46.38	48.61	0.61	n.s.
Involve.	50.54	49.49	0.18	n.s.	49.25	46.26	1.01	n.s.
Primary	N=52	N=35			N=35	N=24		
Read.	100.02	101.66	0.26	n.s.	101.63	102.04	0.01	n.s.
Arith.	102.67	99.77	0.74	n.s.	98.88	101.83	0.56	n.s.
Sci.	49.92	52.77	1.42	n.s.	48.20	49.52	0.23	n.s.
Z Total	252.62	254.20	0.07	n.s.	248.60	253.42	0.09	n.s.
W-S	52.48	48.71	3.93	.10	50.97	48.62	0.92	n.s.
Org.	52.09	49.46	1.80	n.s.	50.86	48.42	0.44	n.s.
Vpt.	46.40	50.28	2.87	n.s.	48.97	51.25	0.58	n.s.
Stab.	51.13	49.23	1.19	n.s.	50.74	48.71	0.49	n.s.
Involve.	53.60	48.65	7.50	.01	50.43	44.42	6.34	.025

TABLE 5-17. Relationships Between TBWS Scores as Norms and Ratings as Norms, Intermediate and Primary Teachers, Lo MFI Districts.

Tasks and Scales	High-Task Oriented, Rating				High Social-emotional Oriented Rating			
	Hi	Lo	F	P	Hi	Lo	F	P
Intermediate	N=28	N=14			N=34	N=14		
Read.	96.50	99.36	0.04	n.s.	99.58	89.07	1.23	n.s.
Arith.	96.50	97.21	0.02	n.s.	96.85	96.92	0.00	n.s.
Sci.	47.67	50.71	0.83	n.s.	40.88	45.64	0.89	n.s.
Z Total	240.67	243.29	0.73	n.s.	240.32	231.64	0.88	n.s.
W-S	49.22	49.14	0.00	n.s.	47.38	46.29	0.16	n.s.
Org.	52.38	51.28	0.19	n.s.	47.74	20.64	0.75	n.s.
Vpt.	50.50	49.50	0.04	n.s.	50.88	49.07	0.27	n.s.
Stab.	50.16	49.42	0.04	n.s.	48.59	48.93	0.01	n.s.
Involve.	52.05	52.36	0.01	n.s.	46.58	47.71	0.11	n.s.
Primary	N=19	N=17			N=31	N=26		
Read.	96.68	99.35	0.02	n.s.	101.26	94.38	3.28	n.s.
Arith.	105.05	100.38	0.61	n.s.	102.68	99.08	0.80	n.s.
Sci.	49.95	49.82	0.00	n.s.	50.65	52.03	0.32	n.s.
Z Total	251.58	250.06	0.02	n.s.	251.35	245.50	0.67	n.s.
W-S	48.47	51.18	0.76	n.s.	50.58	46.77	1.75	n.s.
Org.	47.32	46.53	0.07	n.s.	48.87	51.19	0.84	n.s.
Vpt.	49.95	49.00	0.14	n.s.	48.42	49.46	0.14	n.s.
Stab.	46.78	50.06	1.12	n.s.	50.32	49.27	0.12	n.s.
Involve.	45.84	49.71	0.88	n.s.	54.35	45.65	6.26	.025

TABLE 5-18. Mean Task and Scale Scores of Intermediate and of Primary Teachers Rated Hi vs Lo, According to Per Pupil Wealth of School District.

Tasks and Scales	Hi Wealth				Lo Wealth			
	Hi	Lo	F	P	Hi	Lo	F	P
Intermediate								
Read.	104.12	94.97	6.82	.025	97.34	94.97	0.39	n.s.
Arith.	101.18	94.46	5.24	.05	98.34	98.59	0.00	n.s.
Sci.	51.39	47.95	2.71	n.s.	49.13	49.38	0.01	n.s.
Z Total	256.69	242.17	10.08	.01	244.93	243.28	0.01	n.s.
W-S	50.39	48.16	0.98	n.s.	48.14	48.00	0.79	n.s.
Org.	51.19	50.05	0.20	n.s.	47.16	50.90	2.32	n.s.
Vpt.	51.51	50.32	0.01	n.s.	50.23	51.07	0.12	n.s.
Stab.	51.82	48.41	2.08	n.s.	47.73	49.10	0.43	n.s.
Involve.	51.71	49.60	0.89	n.s.	48.82	48.96	0.00	n.s.
Primary								
Read.	101.66	99.41	0.57	n.s.	96.75	97.27	0.03	n.s.
Arith.	101.38	99.39	0.54	n.s.	101.93	101.41	0.02	n.s.
Sci.	49.68	50.02	0.03	n.s.	50.29	53.00	0.79	n.s.
Z Total	252.72	248.82	0.58	n.s.	247.15	244.37	0.02	n.s.
W-S	50.75	45.25	9.55	.01	49.60	48.32	0.51	n.s.
Org.	49.83	47.75	1.36	n.s.	50.40	48.61	0.41	n.s.
Vpt.	48.05	53.50	8.64	.01	50.02	48.66	0.41	n.s.
Stab.	50.35	46.14	4.62	.05	50.09	49.32	0.13	n.s.
Involve.	50.98	45.16	6.93	.025	52.89	48.54	5.21	0.5

Table 5-19. Mean Task and Scale Scores for Primary and Intermediate Teachers Pooled, Hi VS Lo Ratings, by the Socio-economic Status of Schools and by Wealth of School District.

Tasks and Scales	Middle class rating			Mixed rating			Working class rating		
	Hi	Lo	F	Hi	Lo	F	Hi	Lo	F
Hi wealth									
Reading	104.06	97.19	2.10	101.67	98.81	0.55	102.90	96.21	3.57
Arithmetic	105.59	99.56	1.90	100.65	95.92	1.93	98.88	97.17	0.32
Science	50.88	48.25	0.67	50.87	51.15	0.01	49.75	48.99	0.71
Z Total	260.53	245.00	3.09	253.19	245.88	1.19	251.52	241.38	2.98
W-S	53.72	42.63	17.28***	49.64	51.61	0.72	49.33	44.84	4.28*
Org.	51.31	46.94	1.92	50.54	51.57	0.22	49.60	47.59	0.79
Viewpoint	46.63	53.63	4.63*	50.61	50.31	0.02	49.25	52.21	1.91
Stability	52.22	43.50	7.44**	49.33	49.65	0.02	51.18	48.13	1.82
Involvement	53.00	40.81	12.84***	50.56	52.42	0.58	50.90	46.10	3.70
Lo Wealth									
Reading	98.00	97.00	0.56	98.17	100.88	0.11	94.82	94.23	0.26
Arithmetic	99.12	101.64	0.38	103.39	102.75	0.01	100.78	97.12	0.99
Science	51.17	53.43	0.34	47.05	47.81	0.04	49.30	51.88	1.04
Z Total	248.46	241.96	1.09	248.61	250.75	0.04	242.40	243.23	0.01
W-S	49.63	46.54	2.05	47.89	44.87	0.77	48.95	49.85	0.14
Org.	49.14	51.54	0.97	48.44	46.50	0.44	48.62	50.00	0.30
Viewpoint	49.73	50.57	0.12	51.44	48.50	0.65	49.30	49.31	0.00
Stability	48.59	51.75	2.72	47.28	45.87	0.14	50.42	48.57	0.00
Involvement	51.46	46.35	4.89*	52.61	47.81	2.06	50.30	51.53	0.26

A rather extraordinary aspect of this reversal of pattern is that principals rated primary and intermediate teachers together, that is, these groups were simultaneously ranked. Thus, in order for a reversal of pattern to occur, the rater would have to operate on a differentiated set of criteria, using one set for intermediate teachers and a different set for primary teachers.

Among low wealth districts no pattern of preferred characteristics occurred, with only Involvement among primary teachers appearing as a criterial characteristic. The fact that this characteristic did appear, however, suggests that it is very widely regarded as normative for primary teachers among principals.

For the third analysis, the division between high and low wealth was maintained, and within each wealth cell, schools were separated according to socio-economic status, middle class, mixed and working class, and primary and intermediate teachers pooled. The groups were then divided according to whether the teachers in them were rated high or low by their principal. The results, as shown in Table 5-19, seem to identify middle class schools in high wealth districts as the locus of norms closely identified with personal-social characteristics, although there are some weak effects for working class schools in high wealth districts. In low wealth districts, there are no significant effects, suggesting that if class status is a variable, it operates only in high wealth districts. To other wise account for the effects shown in Table 5-19, however, an additional analysis was necessary.

For the fourth analysis, the distribution of supervisory organizations over the districts grouped by wealth and MFI was plotted. The distribution may be observed in Table 5-20. In this table, the nature of the organization is shown according to the position or titles of the persons involved in it, and the kind of interaction these persons have with each other, as determined from the statements of principals. An arrow with two heads means that there is reciprocal interaction among the persons involved, while an arrow with one head means that the interaction is somewhat lopsided. The latter relation occurred only in system number 12, in which the supervisors visited teachers on a yearly, bi-yearly and tri-yearly schedule, depending on their level of experience in the system, only for purposes of evaluation. In other systems having supervisory personnel, the supervisors worked directly with teachers, usually on a consulting basis. In all instances in which supervisory personnel other than principals were present, an oral report and discussion between the principal and the supervisor occurred following a supervisory visit to a teacher, while in a few systems, the oral report was followed by a written report. In any event, when supervisory personnel were present in a system, there was interaction between these personnel and the principals.

Table 5-20. Distribution of Supervisory Organizations over School Districts, Wealth by MFI.

System	No. of Schools	
Hi Wealth	Hi MFI	
03	3	Teachers ↔ Principals ↔ Elem. Supervisor
05	3	Teachers ↔ Principals ↔ Elem. Consultants
07	3	Teachers ↔ Principals ↔ Elem. Consultants
10	4	Teachers ↔ Principals
11	3	Teachers ↔ Principals ↔ Director Elem. Ed.
13	7	Teachers ↔ Principals ↔ Elem. Supervisor
Hi Wealth	Lo MFI	
01	3	Teachers ↔ Principals
08	1	Teachers ↔ Supervising Principal
15	1	No Rating (Teachers ↔ Principal)
16	5	Teachers ↔ Principals ↔ Elem. Consultants
Lo Wealth	Hi MFI	
04	2	Teachers ↔ Principals ↔ Elem. Coord.
09	2	Teachers ↔ Principal
14	1	Teachers ↔ Principal
17	2	Teachers ↔ Principal
19	2	Teachers ↔ Principal
Lo Wealth	Lo MFI	
02	1	Teachers ↔ Principal
06	2	Teachers ↔ Principal ↔ Curriculum Dir.
12	3	Teachers ↔ Principal ↔ Elem. Supervisor*
18	1	Teachers ↔ Principal
20	3	Teachers ↔ Principal

*Evaluation only

The most striking feature of the distribution of supervisory organizations across the systems in the sample is that they are concentrated almost wholly in the Type 1 (hi wealth-hi MFI) districts. Because of the disproportionate number of schools in the sample from such districts, it is quite obvious that whenever the high wealth or the hi MFI districts are pooled, the resulting cell or group tends to be highly dominated by systems in which there is a distinct supervisory organization. To put the matter another way, both wealth and MFI tend to be confounded with the presence of a supervisory organization. The hi wealth-hi MFI cell is of course almost completely confounded with the presence of a supervisory organization.

With respect to preceding analyses, the possible effects of the presence of a supervisory organization could most clearly appear at four points: first, in the relationship between hi wealth-hi MFI to norms as shown in Table 5-10; second, in the relationship between wealth and norms, as shown in Table 5-18, third, between wealth, SES and norms, as shown in Table 5-19, and fourth, as an unexplicated effect buried in Tables 5-10 and 5-11, and in Table 5-14 in which the decreases in reading performance among teachers in hi MFI-lo wealth districts relative to those in hi wealth-hi MFI districts are unexplained.

To examine the relationships between the presence of a supervisory organization and the presence of norms as indicated through the ratings of teachers, high and low MFI districts were separated, and within each type of district, systems were separated according to whether they had a supervisory organization, i.e. whether they were with or without supervision. Teachers within each of the resulting cells were then divided according to whether they were rated high or low. Three separate analyses were run with this arrangement. The first was for primary and intermediate teachers pooled, for which the results are shown in Table 5-21. As may be observed in this table, the presence of norms is associated wholly with high MFI districts with supervisory organizations. The norms in question are largely concentrated in the personal-social characteristics of teachers, for which the effects are relatively strong, with some weak effects for arithmetic and Z Total.

In subsequent analyses, primary and intermediate teachers were divided, and the analyses re-run under the arrangement described above. The results indicated that for primary teachers, all significant effects were located in high MFI districts with supervisory organizations, and were wholly concentrated in the personal-social characteristics. Differences between primary teachers rated high versus low are as follows: Warmth-spontaneity, $F=11.36$, p .01; Organization, $F=13.69$, p .001, Viewpoint, $F=9.97$, p .01, Stability, $F=5.16$, p 05, and Involvement, $F= 7.04$, p .025. Teachers rated high were more warm-spontaneous, more organized, more child-centered, more stable and more involved. In addition, there was a weak effect for arithmetic, but no effects for reading or science, although the mean differences were in the expected direction.

Table 5-21. Mean Task and Scale Scores for Teachers Rated Hi VS Lo, Primary and Intermediate Teachers Pooled, in Hi and Lo MFI Districts with and without Supervision.

	Hi MFI						Lo MFI					
	With Supervision			Without Supervision			With Supervision			Without Supervision		
Tasks and Scales	Hi N=45	Lo N=27	F	Hi N=64	Lo N=56	F	Hi N=63	Lo N=42	F	Hi N=46	Lo N=34	F
Read	104.18	100.22	3.05	101.86	100.80	0.14	96.94	93.43	1.38	97.30	94.79	0.45
Arith.	103.19	98.70	4.25*	100.92	102.34	0.27	98.75	96.26	0.75	100.35	100.65	0.00
Sci.	50.44	50.92	0.09	51.23	50.86	0.38	48.57	48.62	0.00	51.50	51.73	0.01
Z Total	257.81	249.84	4.25*	254.02	254.18	0.00	242.67	238.31	0.62	249.15	247.18	0.10
W-S	52.11	48.16	6.99***	50.70	49.29	0.74	48.49	48.57	0,02	48.98	48.03	0.16
Org.	50.91	46.83	7.26*	49.01	51.69	1.65	49.46	50.69	0.43	46.70	48.35	0.55
Vpt.	46.09	49.26	4.68*	52.00	51.73	0.02	48.95	48.81	0.00	51.09	50.21	0.15
Stab.	51.05	47.34	6.83**	50.06	50.52	0.06	49.02	50.40	0.47	49.37	48.17	0.30
Involve.	51.08	46.34	9.45***	51.55	47.95	5.12*	47.91	45.95	0.81	51.76	50.06	0.31

Among intermediate teachers, there were no effects in low MFI districts, with or without supervision. In high MFI districts, the effects in districts with and without supervision were similar. In the districts with supervision, teachers rated high were significantly higher in reading (F=4.39, p 05) and Z Total (F=6.43, p 025), with a very weak effect (F=2.88) for arithmetic while in districts without supervision, there were no effects for arithmetic (F=.12), but there were weak effects for reading (F=5.75, p 05) and Z Total (F=4.04, p 05). This particular arrangement of results suggests that a norm for reading is a function of MFI rather than supervision among intermediate teachers.

Returning to Table 5-21 in relation to the results discussed above, one may observe that the significant effect for arithmetic comes about by pooling two weak effects, one from primary and one from intermediate teachers, that the very weak effect for reading occurs because the non-significant results for primary teachers dominates the stronger effect for intermediate teachers, while the effects for the personal-social variables from the Characteristics Schedule are for the most part weakened effects created by pooling the non-significant effects for intermediate teachers with the strong effects for primary teachers.

To understand the effect of supervision on the results shown in Table 5-19, namely, that norms for the personal-social characteristics are identified largely with middle class schools in high wealth districts, it is important to re-examine Table 5-4. As may be noted in the latter table, the pooling of schools across the wealth dimension while maintaining a separation for socio-economic status results in eight of the ten middle class schools falling within hi wealth-hi MFI districts which have supervisory organizations (there was no rating for the one school in system 15; if it were included the ratio would be 8/11). This ratio decreased, however, in the remaining two groups. For mixed schools it is 6/5 and for working class schools, 6/6. Moreover, there are disproportionalities in the number of teachers contributed by systems with supervisory organizations across the three levels of school socio-economic status, with the schools of mixed composition in systems with supervisory organizations tending to contribute proportionally fewer teachers than was true in either the middle class of the working class schools. Thus, what appears to be an effect for middle class schools turns out to be an effect associated with the presence of a supervisory organization.

With respect to the absence of effects in middle class schools in low wealth districts, it should be noted in Table 5-4 that none of these schools are associated with districts having a fully functioning supervisory organization, i.e. an organization in which there is interaction between the teachers and the supervisory staff.

The final set of analyses bearing directly on the effects of a supervisory organization was run wholly within high MFI districts, with primary and intermediate teachers separated and districts divided according to whether or not they supported a supervisory organization. Within districts, teachers were separated according to whether they had been in the system 0-5 or six plus years. As may be noted in Table 5-21a, in which the results are shown, there is significant slippage in the reading, science and Z Total performancesof primary teachers in high MFI districts without supervision, but no slippage at all, indeed a slight increase in these scores among teachers in high MFI districts with supervision. In addition, there is a substantial drop in reading scores between the 0-5 and the six plus groups of intermediate teachers in the high MFI without supervision districts.

The particular arrangement of results shown above, when observed in conjunction with the results in Tables 5-10 and 5-11, and when considered against the fact that the scores of both new and transfer teachers in all high MFI districts are quite similar, distincly suggests that the slippage in task performance in hi MFI-lo wealth districts is attributable to the absence of a supervisory organization in these systems.

Relationships Among Wealth, Median Education Level and Ratings, with School Districts as Observations. The final analysis bearing on norms was conducted primarily as a cross-check of the analyses using teachers as observations, and MFI rather than MEL as the index of community socio-economic status. For the analysis, systems were dichotomized by wealth and by MEL. These dichotomies produce equal frequencies of systems within cells if all systems can be used. System 15, however, did not cooperate in doing the rating. To correct for this difficulty, teachers within cells were divided according to whether they were rated high or low. Within the hi wealth-lo MEL cell, the means of each rating group were then calculated, and the difference found. To obtain a mean for each dependent variable for each rating group in system 15, the mean of its teachers (undivided since there was no rating) was found, then the probable level of the group rated high versus those rated low in this district was estimated to be equal to the mean difference between the groups rated high and the groups rated low in systems drawn from the same cell. This procedure maintains an accurate cell mean for each dependent variable, but underestimates the true within-cell variance since one observation is not permitted to deviate from the mean. In turn, one degree of freedom could be interpreted as being lost, but practically speaking, the loss of one degree of freedom out of 32 is of little consequence.

A second aspect of this analysis may not be apparent, and bears discussion. Under the arrangement used (a 2 x 2 x 2 ANOVA) the estimates of the means for the effects for wealth and the effects for MEL are obtained within the analysis by pooling the means for the rating groups. If the

Table 5-21a. Mean Task and Scale Scores for Primary and for Intermediate Teachers 0-5 and 6 plus Years of Experience in Hi MFI Districts With and Without Supervisory Organizations.

	Primary						Intermediate					
	With supervision		F	Without supervision		F	With supervision		F	Without supervision		F
Tasks and Scales	0-5	6+		0-5	6+		0-5	6+		0-5	6+	
Reading	103.68	102.02	0.29	102.92	94.83	6.67*	103.27	100.40	0.71	108.22	100.92	2.44
Arithmetic	98.74	103.85	3.07	100.48	98.47	0.26	102.20	100.67	0.21	105.41	102.67	0.52
Science	51.43	49.89	0.46	53.13	46.58	6.28*	50.68	51.37	0.09	53.41	51.90	0.26
Z total	253.75	255.77	0.15	256.55	239.84	7.03*	256.15	252.43	0.39	267.41	255.52	2.77
W-S	50.01	50.40	0.04	50.87	49.47	0.50	52.93	47.63	4.59*	52.07	50.28	0.91
Org.	50.64	48.11	1.82	50.61	50.92	0.02	47.97	48.57	0.05	49.18	52.19	1.39
Viewpoint	48.42	47.11	0.40	49.61	53.00	1.65	46.36	47.80	0.46	51.67	50.52	0.14
Stability	49.98	49.79	0.01	50.45	50.11	0.02	49.97	46.53	1.90	51.48	52.00	0.03
Involvement	48.09	50.71	0.97	44.84	53.39	22.51**	46.29	50.00	2.20	47.18	53.71	4.75*

*p = 05

**p = .01

means for each rating group in each system are undergirded by equal number of teachers, the estimates of the means for each rating group might be interpreted as unbiased. In the present data, however, there were both differences in numbers of teachers within each system and in addition, a rating bias among principals, such that there were invariably more teachers rated high than low, even after teachers 0-5 years of experience were removed. Thus, the present analysis treats each rating group as if it were equally weighted, when it is not in fact, and thereby produces a somewhat different estimate of systemic means for the wealth and MEL main effects and interaction than would be the case in a weighted means analysis or in an unweighted means analysis in which ratings were not used. Because of this feature, congruence between the effects in the weighted and the unweighted analysis suggests relatively great robustness of effects, i.e. they are consistent in spite of differences in the method by which estimates of systemic, and within system types, means are made.

For intermediate teachers, the results showed a main effect for MEL, a main effect for rating, and a wealth by MEL interaction for reading, all significant beyond p=.01. There were no significant effects for arithmetic, while in science there was a main effect for MEL, but no other effects, nor were there any effects for any of the personal-social characteristics. The most revealing set of effects lay in Z Total, however. The cell means, F ratios and probabilities are shown in Table 5-22. In this table the most interesting result lies in the second order interaction, wealth by MEL by rating. This interaction issues from a reversal in the ratings in the lo wealth-lo MEL cell. To appraise the effects of this interaction in previous analyses it is important to recall that when districts were dichotomized by MEL rather than MFI, there is an interchange of three systems within the high wealth category, two systems with supervisory organizations (11,07) migrate from the hi wealth-hi MFI cell to the hi wealth-lo MFI cell, while one system, without a supervisory organization (01) migrates from hi wealth-lo MFI to hi wealth-hi MEL. To put the matter another way, classification by MEL rather than MFI re-distributes the systems with supervisory organizations, dispersing them about equally over the two MEL cells within the hi wealth category. The general effect of this dispersion is to strengthen norms in the hi wealth-lo MEL cell, and also to strengthen the effects for reading in that cell, while weakening the effect for reading in the hi wealth-hi MEL cell.

In considering these effects in relation to preceding analyses, it is apparent in Table 5-22 that whenever a conjunction between the two hi MEL cells, the two hi MFI cells, or the two hi wealth cells occurs in an analysis, there is a strong tendency for differences on Z Total, always in conjunction with reading, to appear between those rated high and those rated low. This result does appear not only because there is a positive difference between the rating groups in these cells, but also because there is a reversal or a negative difference in the lo wealth-lo MEL cell which washes out any positive difference which may occur in the particular cell with which it is conjoined in a particular analysis.

Table 5-22. Effects for Z Total Among Intermediate Teachers, and for W-S and Involvement for Primary Teachers, Wealth X Education Level X Rating.

Educ. Level	rating	Z Total Intermediate Wealth		W-S Primary Wealth		Involvement Primary Wealth	
		Hi	Lo	Hi	Lo	Hi	Lo
Hi	Hi	257.80	265.10	49.57	50.47	54.47	52.66
	Lo	242.44	246.60	45.41	47.38	48.75	49.19
Lo	Hi	252.41	227.89	51.93	47.99	49.19	51.40
	Lo	234.62	236.19	42.30	47.12	41.03	45.94

Effect	F	P	Effect	F	P		F	P
Wealth	1	n.s.	Wealth	1	n.s.	Wealth	1	n.s.
Educ. level	19.12	.001	Educ. level	1	n.s.	Educ. level	7.15	.025
Rating	9.72	.01	Rating	7.41	.025	Rating	12.11	.01
W x E.L.	6.11	.05	W x E.L.	1	n.s.	W x E.L.	1.70	n.s.
W x R	2.72	n.s.	W x R	2.64	n.s.	W x R	1	n.s.
E.L. x R	3.07	n.s.	E.L. x R	1	n.s.	E.L. x R	1	n.s.
W x E.L. x R	4.41	.05	W x E.L. x R	1.39	n.s.	W x E.L. x R	1	n.s.

In observing the significant main effect for rating in Table 5-22, care should be taken not to interpret this effect independently of the second order interaction. It would be quite false to conclude from this analysis that, in general, there is a positive norm associated with Z total or total task performance, since such a proposition would be false with respect to lo wealth-lo MEL districts.

Among primary teachers, significant results appeared only on Warmth-spontaneity and Involvement. In each instance, only the main effect for rating was significant, as may be observed in Table 5-22. In examining the effect for rating on Warmth-spontaneity, it may be noted that the effect arises primarily in the high wealth category, with only a slight contribution from the lo wealth-hi MEL cell, but a large contribution from the hi wealth-lo MEL cell. This arrangement produces an interaction wealth X rating which, while not significant, helps clarify the relationships among the various types of systems. In addition one may note that had the analysis been done with a dichotomy on MFI rather than MEL, the greater difference would have appeared in the hi wealth-hi MFI cell since systems with supervisory organizations would have been concentrated in that cell rather than dispersed over the wealth cells as they are in the present analysis.

The effects for the rating on Involvement are relatively homogeneous across the cells shown in Table 5-22, a result congruent with those for Involvement in preceding tables.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

In appraising the results of the sample survey portions of the project, four matters need to be held in mind: 1) the principal dimensions associated with systemic and community variables, 2) the nature of the measures of teacher characteristics, 3) the empirical relationships between these two for primary and intermediate teachers, and 4) the models proposed in Chapter 3.

Intermedicate Teachers, Community Characteristics and Task Performance. Beginning with intermediate teachers, the data consistently show that differences in task performance between teachers in high versus low MFI districts are at their greatest among teachers who have been in these districts five years or less, with decreases in differences as experience in each type of district increases. This pattern of differences suggests that differential selection occurs at a relatively early stage on the task performance characteristics of intermediate teachers, with the high MFI districts consistently holding an advantage. According to the differential attraction model, a difference of the kind in question might be attributable either to the economically controlled policies of the district, or to differential community attractiveness, or possibly, to both. In the present instance, there are three economically controlled policies which could function as an identifiable attraction to teachers: salary incentive, salary limit, and pupil teacher ratio. These variables, however, are identified with the per pupil wealth dimension of communities rather than the socio-economic dimension. Thus, unless there are unidentified economically controlled policies which are a function of either the median family income of the district or its median education level, the best hypothesis is that differences between high and low MFI districts in task performance are in some way associated with differential community attractiveness for teachers.

The differential attractiveness of a community for teachers may be broken into several parts. First, high MFI districts, and especially districts having both hi MFI and hi MEL may be viewed as having more children of middle class origin, relative to the state as a whole, than districts that are low in MFI, or both low in MFI and low in MEL. Thus the probability that a teacher will work with middle class children and middle class parents is higher in such districts. To the degree that teachers in the early and middle 1960's preferred to work with children and parents of middle class origin, communities high on the socio-economic dimension held an advantage in attracting such teachers.

Second, all but one of the high MFI districts in the sample are either themselves industrialized or else lie within easy commuting distance of industrialized areas. As noted in Chapter 3, this factor is important if teaching is considered secondary employment for the married female, since the presence of primary employment for the husband in adjacent industrial or commercial enterprises or supportive services may be viewed as an important condition of whether the teacher appears in the pool from which the high MFI districts can draw their employees.

A third factor, not fully independent of the other two, lies in the competitive position of some high MFI or high MEL districts relative to the districts surrounding them. For example, systems 09, 13, 14 and 19 are all adjacent to SMSA core cities and offer the teacher an escape from the city into a smaller, more middle class, and probably less bureaucratic school system.

The three factors above may be viewed as providing the pool from which the high MFI district may draw teachers. The fact that such a pool might exist, however, does not fully account for differences between high and low MFI or MEL districts. In order to take advantage of an existing pool of teachers who can be employed, it is apparent that the system must by some means select teachers from it. In the differential attraction model, it is assumed that the system has a developed norm or set of norms for the type of teacher administrative personnel believe is successful in the district, and that teachers whose behavior is congruent with the norms are selected.

The extent to which a norm for task performance in the high MFI or high MEL districts appears in the data from the study depends both on the particular aspect of performance observed, and on the controls exercised on the length of time the teacher has been in the system. In both high MFI and high MEL districts, the strongest norm, as judged from the differences in performance between teachers rated high and those rated low, lies in skill in teaching reading. There is, however, sufficient consistency across the total set of tasks for Z Total, a measure of total task performance strength, to appear as consistently associated with the ratings.

The statistical strength of the relationship between ratings and task performance depends both on whether teachers new to the system were excluded from the analysis and on whether the analysis was done with weighted or unweighted means. When the analysis was done with unweighted means and new teachers are excluded, the results are quite strong (Table 5-22), when it is done with weighted means and new teachers are included, the effect appears only in the hi wealth-hi MFI cell (Table 5-10).

A final aspect of the differential attraction model lies in a prediction by the system that the subsequent behavior of the teacher will be congruent with the norm. The data shown in Table 5-14 suggest that high MFI districts do tend to select slightly superior new teachers with respect to task performance. The effect is a weak one, however, and the differences between high and low MFI districts clearly only begin to arise as the teacher enters the system, cumulating more strongly under the selection of transferring teachers who are clearly superior in task performance.

Broadly interpreted, the results of the study with respect to intermediate teachers fit rather well into the differential attraction model. There are certain details of these results, however, which suggest that the behavior modification model, or a modified form of it, is also relevant.

As may be noted in Table 5-7, there is substantial shrinkage in the difference between intermediate teachers in high and low MFI districts in reading performance between the 0-5 years of service group and the six plus years group. If the source of this difference is traced down, as it is in Tables 5-10 and 5-11, it may be observed that the significant effects are concentrated in the lo wealth-hi MFI districts, in which task performance in teaching reading falls from an initially quite high level toward the level characteristics of the low MFI districts. The result is that among teachers of six or more years of service to their districts, the difference between high and low MFI districts is supported almost entirely by the higher mean performance of teachers in hi wealth-hi MFI districts. In addition, it may be observed that it is in the latter districts that both the strongest norms with respect to skill in teaching reading, and a supervisory organization which supports the supervisory activity of the principal, exist.

This particular arrangement of results suggests first, that the presence of a supervisory organization has the effect of maintaining the skill of intermediate teachers in teaching reading, and second, that there is a distinct tendency for stronger norms to appear in systems with supervisory organizations than in systems without such organizations. Considering the latter point in greater detail, it appears to be the case that a supervisory organization is not merely a delivery mechanism through which norms existing in a system are applied, but that the presence of the supervisory organization influences the degree to which the norms appear in the system, when the norms are judged from the differences between the teacher groups rated high and low. While the means by which a supervisory organization could influence the norms of a system, as operationally defined in this study, were not very fully explored, it may be noted (Table 5-19) that in systems in which supervisory organizations are present, there is interaction between the principal and the supervisory staff. Such interaction might influence the presence of norms directly by means of the interaction, i.e. principals and supervisory clarify their own norms through discussing the performance of individual teachers, or it could occur on wholly a procedural basis, i.e. supervisory appraisals based on direct classroom observation and consultation with the teacher may simply be more valid and reliable than those made by principals. It is to be remembered that in every instance in which there was a supervisory organization, a member of the supervisory staff reported to the principal, and principals may simply have used these reports in ranking the success of their teachers for the purposes of the project.

Primary Teachers, Community Characteristics and Task Performance. If it is true, as suggested above, that the differential attraction model, and to a lesser degree, the behavior modification model apply among intermediate teachers with respect to task performance, there is every reason to think that these models might also apply to primary teachers. But they do not in fact apply. With one exception, there are no significant results among all the analyses of task performance conducted using primary teachers. The exception is that reading performance shrinks significantly among primary teachers in both lo wealth-hi MFI districts and lo wealth-lo MFI districts (Tables 5-11 and 5-13), again districts in which supervisory organizations are least present.

The most parsimonious explaination of the general failure to obtain significant effects among primary teachers is that the tasks used to sample performance do not sample the relevant portions of the task universe in teaching reading, and either do not sample the relevant portions of the task universe in teaching arithmetic and science, or else signify that these universes are not significant to the work of primary teachers. In short, the problem-tasks used appear not to be occupationally relevant for primary teachers. The fact that primary teachers in low wealth districts show a downward shift in reading performance, while teachers in high wealth districts do not change significantly in reading performance, suggests that the effects of supervision in the high wealth districts may be to maintain the skills in teaching among primary as well as intermediate teachers. At the same time, there is no evidence that task performance in teaching reading is in any way valued for primary teachers in any type of school district.

To explain this state of affairs, one might quite reasonably assume that supervisory activity, which not only includes room visitation and consultation, but inservice activities of various kinds, tends to maintain a wide range of skills in teaching reading, both among primary and intermediate teachers. Among the total set of skills maintained, however, some are more relevant to intermediate teachers, and some more relevant to primary teachers. In constructing the reading tasks, as discussed in Appendix 1, the assumption was made by the writer that the analytic-diagnostic type of task relevant to intermediate teachers would also be relevant for primary teachers if the content for primary teachers were simply shifted toward grades 1-3. As matters turned out, however, no evidence could be generated in any portion of the project to firmly substantiate this assumption, and the assumption might be regarded as false. Under this interpretation, in order to have shown results in reading among primary teachers similar to those for intermediate teachers, one would have to have constructed a quite different kind of task for primary teachers than those employed in the study.

The same arguments might be applied to the arithmetic and science domains for primary teachers. In the instance of science, however, there is perhaps reason to question whether, in comtemporary teaching practice at the primary level, an occupationally relevant sample of tasks could be drawn from this domain. Until quite recently instruction in science has received very little emphasis in the primary grades, and it is perhaps questionable whether any school system would hold skill in teaching science as strongly normative in these grades. Indeed, there is no evidence in the data from this study that ability to perform the type of science task used is normative for primary or intermediate teachers in any type of school district.

Task performance in arithmetic for primary teachers seems to lie between science and reading in its significance to the primary teacher. While it may be possible to construct occupationally relevant arithmetic tasks for primary teachers, whether strong norms would appear with respect to them in any school system remains, at least in the mind of the writer, an open question.

Primary Teachers, Community Characteristics and Personal-social Characteristics. Among primary teachers, the general pattern of scores suggested no differences in personal-social characteristics between systems of different types, although there was one analysis in which Involvement appeared as significantly related to the median education level of the community (Table 22). In spite of this general pattern, examination of the differences between teachers 0-5 and six or more years of service to their system revealed that there were downward shifts in scores on Warmth-spontaneity and Stability, and an upward shift on Viewpoint (toward greater subject-centeredness) in low MFI districts, but an upward shift in Involvement scores in high MFI districts, with no downward shift on other characteristics. The strength of the shift was sufficiently great in Warmth-spontaneity to produce a significant difference between high and low MFI districts between teachers six or more years of service to their system, and on Involvement apparently strong enough to support the over-all difference cited above. Generally, however, difference between system types in Involvement was not a strong effect.

Unlike the pattern of results for intermediate teachers in task performance, the pattern for primary teachers suggests that there is no differential selectivity occuring between systems during the initial or early stages of employment, but it does suggest that either selectivity or differential retention or holding power occurs between systems after a period of employment. Moreover, the differential follows the same community variable, median family income, as was the case among intermediate teachers in task performance.

In accounting for the results among primary teachers, it is first important to observe that they could not reasonably have occured under the behavior modification model. If this model applied, differences in score levels should have appeared between the hi wealth-hi MFI districts and lo wealth-hi MFI districts, since supervisory organizations are associated with the former but not the latter, and since it is quite clear in Table 5-20 that the norms for personal-social characteristics are heavily concentrated in districts with supervisory organizations. Thus, if either of the two models suggested in Chapter 3 apply, the most relevant should be the differential attraction model.

The apparent initial weakness in applying the differential attraction model is that there is no evidence of differential attraction between system types with respect to the personal-social characteristics of primary teachers recent to their systems. Score levels in this group are quite similar from one type of district to the next. Failure to show differences among the groups recent to their systems, however, does not automatically eliminate the model. In order for the model to operate, there must not only be an available pool of teachers to employ, but selection must be made from this pool. Thus, the failure for differences to appear may indicate difficulties in selection procedures.

Quite clearly, in order for systems to initially select teachers on the criterion of personal-social characteristics of the type employed in the study, some way of identifying these characteristics prior to employment would be necessary. As pointed out in Chapter 1, however, the characteristics used in the project were primarily induced from the observation of the classroo behavior of teachers. To suppose that these characteristics could be reliably identified prior to the employment of a teacher entails that the employer has some means of obtaining evidence with respect to the characteristics. Such evidence would be unlikely to be found in the traditional materials used to select teachers, i.e. school records, letters of recommendation, and interviews. In short, one may suppose that in predicting congruence between the norms of the system and subsequent teacher classroom behavior, most systems error very widely in their prediction. The prediction may be viewed as no better than chance, and indeed the result is a random distribution of personal-social characteristics across systems among primary teachers. Nonetheless, the very fact that the personal-social characteristics employed in the study can be induced from observation means that as the experience of a teacher increases in a system, and as the number of observations of that behavior also therefore increases, a reasonably accurate portrait of her characteristics can be formed by supervisory personnel or by principals. Thus, given the nature of the personal-social characteristics employed in the study, there are reasonable grounds for supposing that the selection process among primary teachers might be gradual rather than abrupt.

If one assumes that gradual selection processes do operate in high MFI districts, it seems evident that they do not operate in precisely the same way in hi wealth-hi MFI districts and lo wealth-hi MFI districts. In hi wealth-hi MFI (or high MFI districts with supervision) there are consistent indications that all, or virtually all, of the personal-social characteristics or primary teachers are held as normative. In low wealth-hi MFI districts, on the other hand, only Involvement recurs as a norm. Moreover, upward shifts in Involvement scores occur wholly within the lo wealth-hi MFI districts, in which, as exemplified in Table 5-21a, the effect is very strong. Because Involvement scores are positively associated with scores on the other personal-social scales, selection of primary teachers on the criterion of Involvement alone would have the effect of maintaining approximately equal score levels on the other personal-social scales between teachers 0-5 and six plus years of experience in these districts. This argument is not operable for hi wealth-hi MFI districts, however. In these districts there is no evidence that selectivity occurs on a specific personal-social variable. Rather, the maintanence of similar score levels between teachers 0-5 and six plus years of experience in their district must be interpreted to have occured by means of selection across the range of personal-social variables for which there are norms in the district. These differences in selection processes, it might be noted, align themselves with whether or not there is a supervisory organization in the system. When the principal performs the supervisory function alone, Involvement appears to be the over-riding criterion, but when a supervisory organization is present, the criteria appear to be broadened to include the other personal-social characteristics. Again, these inferred differences in selection process between the two types of high MFI districts appear to be associated with the kind of evidence available to the person performing the supervisory function. Principals are perhaps in the best position to know which teachers are committed to their work in the sense that they know who stays after school, who attends meetings, who takes work home, and who prepares most conscientiously for instruction. Supervisors, on the other hand, are perhaps in the better position to observe on-going classroom behaviors and therefore make judgements concerning the broader spectrum of classroom personality factors.

The argument that high MFI districts select primary teachers in such a way as to maintain similar levels of characteristics between teachers 0-5 and six plus years of service to the system may be viewed as a viable, but not completely satisfactory argument. It is true, for example, that Involvement is generally held as normative among systems (Tables 5-22 and 5-18). Thus, if systems select on this norm, there should be no type of system in which differences develop between the more and the less experienced teachers in their system. Such differences do occur in low MFI districts, however, and especially in hi wealth-lo MFI districts, where the differences between the two groups on Warmth-spontaneity is quite marked, with associated differences in Organization and Viewpoint. There are no differences, however, in Involvement scores and only weak differences in Stability. This particular pattern of results suggests that some other factor or set of factors may be operating.

A cue to these factors may be found in Tables 15 and 15a. In Table 15 it may be noted that there is a substantial discrepancy in Warmth-spontaneity score levels between primary teachers who are new to the low MFI districts and those who have transferred to the system in the past, with the transfers showing a score level about one Z score standard deviation below those new to the system. Moreover, this discrepancy is an important one in terms of the number of teachers involved since more than half the primary teachers in low MFI districts have transferred to them from other systems, as shown in Table 5-15a. Keeping in mind that selectivity would have operated within the transfer group since many of the teachers in the group have been in the system in which they are teaching for five years or more, it nonetheless appears to be the case that the pool of teachers from which the low MFI district draws its transfer teachers is very different from the pool from which it draws its new teachers. To put the matter a bit differently, if primary teachers are drawn at random from a pool of potentially employable teachers, the pool of experienced teachers from which low MFI districts draw their transferring group is distinctly low in Warmth-spontaneity. Moreover, there is some tendency for this group to be lower on Organization, higher on Viewpoint, and lower on Stability. In sum, the situation is one which suggests that low MFI communities lie at a disadvantage in attracting a pool of experienced teachers at or above the state mean in personal-social characteristics, hence, that the differential attraction hypothesis is relevant to the slippage shown in low MFI districts in Warmth-spontaneity, Organization and Viewpoint, if not also Stability.

A final factor to be considered also may be inferred from Table 15. Across most of the personal-social variables there is some discrepancy between the scores of teachers new to their sytem and those always in the system. Assuming that those new to the system are not different from a random sample of teachers who in the past were new to the system, but are now in the "always in" group, it is quite clear that there is some differential attrition in the low MFI districts.

Over-all, the downward shifts in score levels on the personal-social characteristics of primary teachers in low MFI districts seem to come about through a series of slight disadvantages suffered in these systems. They seem unable to selectively retain the higher scoring new teachers, with some indication that those scoring higher in fact leave systems of this type. In addition low MFI districts appear to draw from a pool of experienced transfer teachers among whom the level of scores on Warmth-spontaneity and associated variables is substantially lower than in the pool from which they draw new teachers, and somewhat lower than the pool from which high MFI districts draw their transferring teachers. One indeed may suspect that the high MFI districts are taking the higher scoring primary teachers away from the low MFI districts after they have had a few years of experience.

Intermediate Teachers, Community Characteristics and Personal-social Characteristics. Among intermediate teachers, the results relating community characteristics to teacher characteristics are difficult to interpret. First, there is little indication that any of the personal-social characteristics are consistently normative for intermediate teachers in any type of district. Only one variable, Organization, appears as a norm, and then only in low wealth-hi MFI districts (Table 5-16). The same effect recurs in high MFI districts among social-emotionally oriented principals, but not among task oriented principals (Table 5-11). The effect also recurs in high districts among social-emotionally oriented principals, but not among task oriented principals (Table 5-16). The direction of the effect is for teachers rated high to be lower in Organization than those rated low. It may be noted that this effect is the reverse of that for primary teachers in hi wealth-hi MFI districts with supervision. Similar, but less noticable reversals occur with respect to Viewpoint and Stability in Tables 5-10 through 5-13. The situation appears to be one in which principals in certain types of schools or school districts place negative value on the same level of a teacher characteristic that other principals value positively. There is, as it were, a true interaction. The problem is that the particular variables entering this interaction could not be located in analyzing the data, and significant or strong effects could therefore not be shown, even though there is some reason to believe that such effects exist. Indeed, the TBWS was designed precisely to bring out such effects, but proved to be ineffective in doing so. In any event, there is no indication that either behavior modification or differentail selection could occur in any type of school district on the basis of norms linked to the personal-social characteristics of intermediate teachers.

A second set of effects, somewhat less difficult to interpret, are those shown in Table 5-3, in which teachers in high MFI districts are significantly more Warm-spontaneous than those in low MFI districts, and in which there is an interaction wealth by MFI for Organization and Stability. In all cases the effects are carried primarily by the very low scores for intermediate teachers in lo wealth-lo MFI districts. It is notable that intermediate teachers in the latter districts tend to be generally weak, with task as well as personal-social scores substantially below the state average.

In spite of the apparent advantage of high MFI districts in having more Warm-spontaneous teachers, examination of this effect according to years of service the teacher has given to the system, as shown in Tables 5-10 through 5-13 clearly indicates that slippage in Warm-spontaneity occurs in all district types except lo wealth-lo MFI, in which scores are uniformly low. In accounting for this slippage, Table 5-15 is quite helpful; it may be observed there that all transferring intermediate teachers are relatively low in Warmth-spontaneity. In high MFI districts, however, both teachers new to the system and teachers always in the system are high in Warmth-spontaneity. Thus, the slippage in Warmth-spontaneity scores in high MFI districts appears to be primarily traceable to teachers who have transferred to them and have more than five years of service in the system.

Returning to the earlier interpretation that high MFI districts tend to select intermediate teachers high in task performance, it appears to be the case that in selecting teachers on this set of characteristics, the system either cannot or does not control the variability on the orthogonal set of characteristics, i.e. those of personal-social nature. Thus, selecting on the basis of one set of characteristics, systems seem to incur a disadvantage on one or another of the characteristics in the orthogonal set to which, perhaps, no attention is given at the time selection occurs.

Model Revisions and Conclusions. Both the behavior modification-selection model and the differential attraction model presented in Chapter 3 assume that differences in teacher characteristics between school districts flow from differences in the communities in which the systems are located. In the study, three orthogonal dimensions, property wealth, socio-economic status and urbanization were found to describe the major lines along which the communities sampled differ. Of these dimensions, only property wealth and socio-economic status were found to be in some way associated with differences in teacher characteristics. Thus, the undifferentiated concept "school and community variables" presented in the models in Chapter 3 can be differentiated into two relevant orthogonal sets of variables, property wealth and socio-economic status. This orthogonal separation is shown in Figure 6-1.

Considering the observed effects coordinate to the two dimensions in question, one may first note that higher property wealth is associated with the presence of supervisory organization or what was earlier called a "delivery mechanism" in school systems, while lower property wealth is associated with the absence of a supervisory organization. In turn, the presence of a supervisory organization was shown to be associated with the presence of system norms for teacher characteristics, as judged from the relationships between the level of characteristics of teachers rated highly successful in their school as opposed to the level of characteristics of teachers rated not highly successful. These norms, however, were not unitary across primary and intermediate teachers; rather, a high level of task performance in teaching reading, and to a lesser degree, a high level of task performance in teaching arithmetic, were normative for intermediate teachers, but not for primary teachers. Among primary teachers, a high level of Warmth-spontanteity, child-centeredness (low Viewpoint), Stability, and Involvement were normative but these characteristics were not normative for intermediate teachers. Further, task performance variables and personal-social variables (Warmth-spontaneity, Stability etc.) were shown to be orthogonal among both primary and intermediate teachers, thus making it possible for orthogonal norms to appear for these two groups. The orthogonal norms for primary and intermediate teachers, as associated with the presence of a supervisory organization, are also shown in Figure 6-1.

As well as being coordinate to the property wealth dimension of communities, effects relative to norms were also observed for the socio-economic dimension. Again, these norms were orthogonal for primary and intermediate teachers, but they were substantially weaker than those associated with the wealth dimension. Among intermediate teachers the high task performance in reading was normative, while among primary teachers, a high level of Involvement was normative.

The fact that the effects for norms were weak on the socio-economic dimension of communities, but very strong, especially for primary teachers, in systems in which supervisory organizations were present indicates that the relationships between systemic norms and the supervisory organization suggested in the behavior modification-teacher selection model is probably in error. Supervisory organizations appear not to "deliver" or merely apply or utilize norms present in the system, but also appear to generate or at least clarify the norms. High wealth districts with supervisory organizations produce a very clear profile of the type and level of characteristics valued in these systems at both the primary and the intermediate levels.

In connection with the differential attraction model presented in Chapter 3, the possibility was discussed that elementary teachers might be differentially attracted to either school systems or to communities. If they were attracted to school systems, they should be attracted to high wealth systems where there are salary and pupil-teacher ratio advantages. If they were attracted to communities, they should be attracted to communities with higher socio-economic status. On the basis of the results of the study, the inference was made that socio-economic dimension of communities was the prevailing factor. This inference was based on two observations among the results for intermediate teachers. First, differences between communities in the task performance of intermediate teachers were associated wholly with the socio-economic dimension. These differences in turn appear to arise through a series of slight advantages enjoyed in high socio-economic communities in selecting both beginning and transferring (or mobile) teachers. These selection advantages can come about in only two ways. The system either deliberately selects the higher performaning teachers from a large available pool, or else it randomly selects from a pool (small or large) which is homogemeous and high in task performance. In short, the pool must either be large (and partly good), or simply good. Either way, the high S-E district wins relative to the low S-E district. Because the pool in question is a wholly inferred entity, it is shown in parens in Figure 6-1.

As may be further observed in Figure 6-1, beginning with the inference that differentiated pools of teachers exist for high and low S-E districts, two additional inferences were made, both of which imply that the pool of teachers from which high S-E districts draw is large and partly good rather than simply good. These inferences are determinant, however, with respect to the results of the study. These results indicated that there are no differences among primary teachers, 0-5 years of service to their system, attributable to the type of system in which they teach. There are sharp differences between systems among intermediate teachers, 0-5 years of service, in task performance, however. In order for these results to occur, high S-E districts could not be drawing their teachers from a homogeneous and superior pool of teachers. If this were the case, there would be differences among the primary teachers according to the type of system in which they teach, which there were not.

The observed effects for primary and intermediate teachers entail that there is a small but consistent selective factor operating in high S-E districts in choosing intermediate teachers high in task performance. This selective factor may be viewed as a slightly better than chance prediction concerning who in the pool of teachers to be employed will be superior in task performance. Among primary teachers, this prediction is not better than chance, resulting in a failure to find differences according to system type among teachers 0-5 years of service to their system.

The possibility that the relationships noted above could occur may be viewed as a function of three factors. First, the differentation of norms for primary and intermediate teachers. Second, the orthogonal relationship between task performance and personal-social variables among both primary and intermediate teachers, and third, the nature of the personal-social and task performance variables themselves, with the personal social characteristics of teachers held to be less predictable on the basis of typically available evidence than are the task performance characteristics.

Up to this point, the model in Figure 6-1 may be viewed as a variation of the differential attraction model presented in Chapter 3. The variation from the original model occurs primarily through the separation of effects for the major dimensions of school communities, and for primary and intermediate teachers. Beyond this point, the behavior modification-selection model is the more applicable of the two models originally presented.

Among intermediate teachers differences in task performance in reading between low and high S-E districts shrink as service to the system passes the fifth year. The only reason that the differences does not disappear altogether lies in the ability of the high wealth-high S-E districts to maintain high task performance levels of their intermediate teachers. The ability to maintain performance in these districts is directly associated with the presence of supervisory organizations in them. Collaterally, it is the absence of supervisory organizations in lo wealth-high S-E districts which is associated with a decline in the task performance in teaching reading as service to the district increases.

This particular set of results is not wholly congruent with the concept of behavior modification introduced in the original model since behavior seems not so much to be modified as maintained. It is for this reason that the relationship is W + in Figure 6-1. Yet, ability to maintain skills against the onslaught of variables which erode them is no mean accomplishment, and relationship between the presence of a supervisory organization and the maintanence of task performance skills should not be regarded as insignificant.

Among primary teachers an effect similar to that discussed immediately above for intermediate teachers also appeared, but in relation to Warmth-spontaneity and Viewpoint. Both of these characteristics remain at relatively high levels in high S-E districts among teachers with more than five years of service to the district, but the levels drop sharply for the comparable group of teachers in low S-E districts. Since the effect occurs in high S-E districts irrespective of whether a supervisory organization is present, the possibility that the behavior is maintained by this mechanism may be discounted. Rather, the central mechanism appears to be one of differential selection, with hi wealth-hi S-E districts selecting on the full set of personal-social characteristics as norms, while the lo wealth-hi S-E districts select primarily on the basis of Involvement. In addition, there was evidence that low S-E districts draw their transferring teachers from a pool with low personal-social characteristic scores, and that the drop in performance in these districts among teachers past the fifth year of service is in part attributable to this factor.

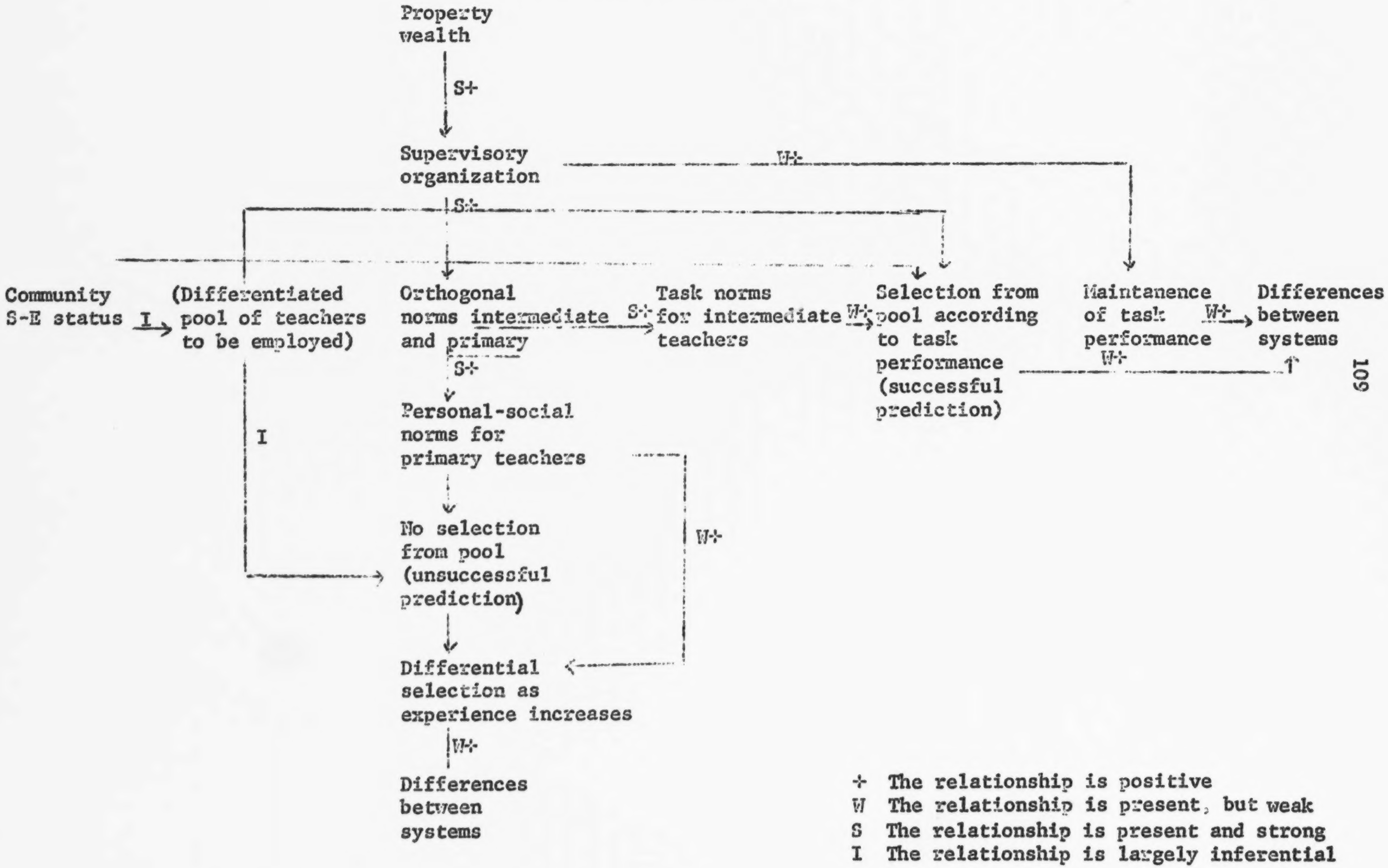
Over-all, one may conclude from the study that differences between high and low S-E districts in the task performance levels of intermediate teachers arise through a slight advantage of the high S-E districts in the market place, together with modest but consistent ability to take advantage of selection opportunities, and ability to maintain performance through supervisory organizations. None of these effects are in themselves very strong. Yet if the accumulated or gross differences between high and low S-E districts in total task performance is translated into the probability that a child in the intermediate grades will be taught by a teacher above the state average in task performance as opposed to below the state average, the results are clear. In high S-E districts the chances are in excess of two to one that a child will have a teacher above average in task performance, while in low S-E districts, the chances are greater than two to one that a child will have a teacher whose task performance is below the average for the state.

There was very little evidence in the study, on the other hand, that there are differences between school districts with respect to the personal-social characteristics or either primary or intermediate teachers, or with respect to the task performance of primary teachers. There was substantial evidence that high levels of Warmth-spontaneity, child-centeredness, emotional stability and involvement in teaching are valued in districts with high S-E status, and especially in those districts within this group in which supervisory organizations are present. Yet, in spite of these apparent values, these systems seem unable to gain a firm advantage in either attracting or developing high levels of these characteristics among the elementary school teachers they employ.

Applications. There are two principal applications of the results of the study. First, there is substantial evidence to suggest that high S-E districts value for primary teachers high levels of the personal-social characteristics measured, but that they are unable to select from the available pool teachers who are high in these characteristics. To increase the congruence between the norms applied and actual characteristics of the teachers employed, utilization of instrument such as the BDT-Characteristics Schedule as a screening device appears to be advisable, at least from the viewpoint of the system. In broader social perspective, however, this move would give high S-E districts even a further advantage in obtaining superior teachers, perhaps with the long range effect of decreasing the probability that children in low S-E districts would encounter primary teachers with high levels of the personal-social characteristics in question.

Second, there is abundant evidence that the lo wealth-lo S-E districts in Indiana suffer a distinct disadvantage in attracting and maintaining intermediate teachers whose task performance skills are at a high level of proficiency. While the long range solution to this problem undoubtedly lies in developing a better industrial-commercial property base in these districts, a reasonable short-term solution might be to provide supervisory organizations for these systems by means of external funds from state or federal sources. These districts are already poor and clearly cannot provide these services from local revenuues without decreasing services in some other area. According to the evidence from the study, the provision of supervisory organizations for these systems would serve to sharpen the norms for teacher behavior and characteristics in these districts, and there is much evidence that these districts are now essentially normless, and in addition, provide the means by which to up-grade the level of skills in the resident teacher population.

FIGURE 6-1. A Modified Teacher Attraction-Behavior Modification Model.


Property wealth
S+
Supervisory organization
W+
S+
Community S-E status
I
(Differentiated pool of teachers to be employed)
Orthogonal norms intermediate and primary
S+
Task norms for intermediate teachers
W+
Selection from pool according to task performance (successful prediction)
Maintanence of task performance
W+
Differences between systems
W+
S+
Personal-social norms for primary teachers
I
No selection from pool (unsuccessful prediction)
W+
Differential selection as experience increases
W+
Differences between systems
+ The relationship is positive
W The relationship is present, but weak
S The relationship is present and strong
I The relationship is largely inferential

References

1. Alpert, Richard and Haber, Ralph N., "Anxiety in Academic Achievement Situations," *Journal of Abnormal and Social Psychology*, 61:207-215, September, 1960.

2. Brown, Rex, *Teacher Skill and Other Variables Associated with Effectiveness in Teaching Reading*, Unpublished Doctoral Dissertation, Indiana University, (in process).

3. Coleman, James, *Equality of Educational Opportunity*, U.S. Department of Health, Education, and Welfare, USOE, OE 38001, 1967.

4. Denny, David A., "Identification of Teacher-Classroom Variables Facilitating Pupil Creative Growth," *Journal of the American Educational Research Association*, 5:365-383, May, 1968.

5. Dreger, Ralph M. and Aiken, Lewis R., The Identification of Number Anxiety in a College Population," *Journal of Educational Psychology*, 48:344-351, October, 1957.

6. Dutton, Wilbur H., "Attitudes of Prospective Teachers Toward Arithmetic," *Elementary School Journal*, 52:84-90, October, 1951.

7. Flanders, Ned A., *Teacher Influence, Pupil Attitudes and Achievement*, Cooperative Research Monograph No. 12, OE 25040, U.S. Department of Health, Education, and Welfare, Office of Education, 1965.

8. Gray, William S., *On Their Own in Reading*, Scott, Foresman, Chicago Rev. ed., 1960.

9. Gross, Edward, *Work and Society*, Crowell, New York, 1958.

10. Guilford, J.P., Merrifield, P.R. and Cox, A.B., "Creative Thinking in Children at the Junior High School Levels," *Reports of the Psychological Laboratory*, 20:1-35, University of Southern California, 1961.

11. James, H.T., Thomas, J.A., and Dyck, Harold J., *Wealth, Expenditure and Decision Making*, Stanford University, School of Education, 1963.

12. Kiesling, Herbert J., "Measuring a Local Government Service: A Study of School Districts in New York State, "*Review of Economic and Statistics*, 356-367, August, 1967.

13. Mills, Belen, *Relationship of Career Attitudes to Stability in Teaching Among Female Elementary School Teachers*, Unpublished Doctoral Dissertation, Indiana University, 1966.

14. Quinn, Edward D., Relationships Among Teacher Characteristics, The Predominant Socio-Economic Class of Students, and Teacher Effectiveness, Unpublished Doctoral Dissertation, Indiana University, 1968.

15. Reisert, John, A Study of Selected Characteristics of Elementary Teachers with Different Preparation Backgrounds, Unpublished Doctoral Dissertation, Indiana University, 1965.

16. Ryans, David G., Teacher Characteristics Schedule, Form E54, 1954.

17. Ryans, David G., "Teacher Behavior, Theory and Research: Implications for Teacher Education," Journal of Teacher Education, 14:274-293, September, 1960.

18. Soar, Robert S., An Integrative Approach to Classroom Learning, Temple University, Philadelphia, 1966.

19. Spache, George D., Diagnostic Reading Scales, California Test Bureau, Monterey, California, 1963.

20. Turner, R.L., Quinn, E.D. and Smith, N.W., Skill in Teaching Assessed on the Criterion of Problem Solving, Three Studies, Bulletin of the School of Education, Indiana University, Vol. 39, No. 1, Bloomington Indiana, January, 1963.

21. Turner, Richard L., Problem Solving Proficiency Among Elementary School Teachers, USOE CRP No. 1262, Indiana University, August, 1964.

22. Turner, Richard L., "Some Predictors of Problems of Beginning Teachers," Elementary School Journal, 67:5, 251-256, February, 1967.

23. Turner, R.L. and Denny, D.A., "Teacher Characteristics, Classroom Behavior and Changes in Pupil Characteristics," Paper American Education Research Association, Chicago, February, 1967.

24. White, Kinnard P., Professional Career Involvement Among Female Elementary School Teachers, Unpublished Doctoral Dissertation, Indiana University, 1964.

25. Yoder, Samuel, Relationships Between Student Teaching Problems and Selected Student Teacher Characteristics, Doctoral Dissertation, Indiana University, 1967.

26. York, Jean, Relationships Between Problems of Beginning Teachers, Their Personal Characteristics and Their Preferences for In-Service Education, Unpublished Doctoral Dissertation, Indiana University, 1967.

APPENDIX I

INSTRUMENT DEVELOPMENT

The instrument development phase of the project was occasioned by the outcomes of an immediately preceding study by the investigator (USOE CRP 1262) in which changes over a two year period in the performance of approximately 200 beginning primary and intermediate teachers in teaching reading and arithmetic, the association between performance and personal-social characteristics, and the relation of each of these to success in different types of school systems were examined. In the study, problem solving performance in teaching reading and arithmetic among both primary and intermediate teachers was assessed by means of "teaching tasks in reading" and "teaching tasks in arithmetic," each containing four problem-tasks designed primarily, although not exclusively, for intermediate teachers. The personal-social characteristics of the teachers were assessed by means of D.G. Ryans' Teacher Characteristics Schedule.

Briefly, the outcomes of this study were as follows:

1. The relationships between the problem-solving characteristics and the personal-social characteristics of beginning elementary teachers are orthogonal.

2. Intermediate teachers teaching in high-wealth school districts increase in performance in teaching arithmetic during the first two years of experience, but intermediate teachers in low-wealth districts do not increase in performance. The increases in performance in high-wealth districts are largely attributable to the prevalence of supervision in these districts.

3. Primary teachers do not increase in performance in teaching arithmetic during the first two years of experience.

4. Both primary and intermediate teachers increase in problem-solving performance in teaching reading during the first two years of experience, with differential increases associated with the amount of supervision received, but only in high-wealth districts.

5. Performance in teaching reading does not predict success in teaching, as viewed by supervisory personnel, among either intermediate or primary teachers, although it is concurrently associated with success at the end of the second year of experience.

6. Performance in teaching arithmetic predicts the success of intermediate teachers in high-wealth districts, but is slightly negatively related ($p<.10>.05$) to success among some intermediate teachers in low wealth districts.

7. Performance in teaching arithmetic is unrelated to success among primary teachers.

8. When factor analyzed alone, the problem-solving tasks in reading and arithmetic produce ambiguous factor structures: for intermediate teachers, the diagnostic tasks in reading and arithmetic load on one factor, while the ordering tasks load on a separate factor; among primary teachers, however, the reading and arithmetic tasks load on separate factors with no separation of diagnostic versus ordering skills.

9. When factored with the problem-solving tasks, the scales in Ryans TCS produce within themselves a single unitary factor.

10. For primary and intermediate teachers pooled, selected TCS scales predict success in teaching among beginning teachers in low-wealth districts.

11. The TCS scales do not predict success for primary and intermediate teachers in high-wealth districts.

12. TCS scales X (friendly, and understanding), Y (business-like), Z (stimulating, imaginative), and Q (relations to school staff) predict beginning teacher problems with discipline, teaching reading, and setting appropriate pupil expectancy. Performance in teaching reading and arithmetic, combined, predicts problems in discipline and setting appropriate expectation.

13. Problem-solving performance in teaching reading does not predict problems with teaching reading.

In addition to these outcomes, a relevant observation made during this study was that to perform the problem-solving tasks and complete the TCS absorbed between four and five hours of teacher time, with about two hours going to the problems and about three hours to the TCS.

The outcomes cited above, together with the observation noted, suggested a distinct set of instrumentation problems. First, there was no indication that the arithmetic tasks utilized were relevant for primary teachers. Because these tasks were developed primarily for intermediate teachers, the inference was made that the tasks were probably at fault in the sense that they sampled a universe of skills irrelevant to the work of primary teachers. Second, there was substantial evidence that the reading task employed did not constitute an adequate sample of the occupationally relevant skills of intermediate teachers, and also appeared to be peripheral to primary teachers. While it is true that the combined arithmetic and reading task score for intermediate teachers did predict success in certain types of school systems, as well as the occurence of discipline and expectancy problems across all systems, the task scores in arithmetic carried the main burden of these results, with the reading tasks carrying a portion of the variance typically significant at about the 10 percent level of confidence.* Thus the reading tasks employed were interpreted to be weak, although not totally irrelevant to teaching in the intermediate grades. Third, the factor structure of the reading tasks, while reasonably interpretable for intermediate teachers was virtually uninterpretable for primary teachers. In combination, these three difficulties indicated that a distinct revision of the tasks in teaching reading would be necessary, that additional arithmetic tasks were needed, and that the tasks for intermediate teachers would have to be separated from those for primary teachers.

A fourth difficulty with the instruments employed centered on the TCS. The basic problem with this schedule lay in the fact that in the beginning teacher population sampled, the intercorrelations among the 10 scales were very great, invariably producing a single general factor under factor analysis. At the same time, in Ryans' earlier studies, sufficient independence across the various scales existed for him to consider each scale a separate "factor." The situation thus appeared to be that with some revision a shorter schedule could be produced, perhaps with the possibility of increasing the independence of some of the scales through differential item selection.

A final difficulty with the instruments, as a group, lay in the length of time required to complete them. The time limit for the assessment of teacher characteristics among current populations of experienced teachers probably lies between two and two and one-half hours if even a semblance of external validity in the study employing the instruments is to be maintained. Beyond this limit, the probability that a system will agree to participate in a study, and the probability that individual teachers within systems will actually complete participation, falls off very rapidly.

In addition to the matters described above, two other types of considerations entered the instrument development phase of the project. First, performance of tasks in teaching reading and arithmetic is unstable in the final year of teacher preparation and the first two years of teaching experience. This instability may by hypothesized to arise through a number of cognitive and personal-social characteristics interacting with particular forms of learning treatments. There is reason to suspect that particular characteristics are involved in these interactions.

---

*The parallel form reliability for both the reading and arithmetic tasks was about .75, thus differences in the strength of the two sets of tasks cannot be attributed to differences in their reliabilities.

A study by White (24) in conjunction with the preceeding project showed that when the degree of career involvement of elementary teachers was scaled, the resulting variable was significantly associated with changes in task performance in teaching and arithmetic in those school systems in which there was an absence of supervision, but not when there was prevalent supervision. This particular set of results suggested that the extent to which the task performance will change in the absence of external motivator-instructors (supervisors) is contingent on the extent to which the beginning teacher is involved in and committed to teaching as a career. Career involvement is not a factor in performance change, however, when an external motivator-instructor is present. Thus, there is interaction between the motivation of the teacher and the type of "treatment" received.

A second teacher characteristic, or set of two characteristics perhaps, suspected to interact with task performance was abduced from two types of evidence. The first was the periodic separation under factor analysis of the tasks requiring diagnosis of pupil difficulties in reading and in arithmetic from those requiring the ordering or organization of materials in these areas. The second was recuring moderate positive correlation between intelligence and task performance. The fact that intelligence related to task performance suggested that performance was associated with a stable intellectual characteristic, while the separation of the two types of tasks under factor analysis suggested that the particular abilities involved might be linked either to induction generally, or to two subordinate abilities, one associated with concept attainment and the other associated with ordering or sequencing concepts or instances of concepts. The reason that both of the latter abilities may fall within a general inductive ability lies in the fact that the ordering or sequencing of either concepts or instances of them suggests the presence of an implicit rule by which concepts or instances may be ordered, a process not unlike induction as concept attainment.

The central problem in ferreting out abilities underlying task performance lies in the great over-mantle of specific skills and achievements which enter into it. For example, to diagnose pupil difficulties in arithmetic one must know a substantial amount about the structure of arithmetic and he must also have developed a complex set of skills associated with the most probable places in this structure to look for errors made by pupils at particular levels of attainment. To solve this problem in the present project, the best initial strategy appeared to be to construct measures of ordering or organizing skills and of concept attainment which were relatively free of specific subject-matter content.

A final teacher characteristic suspected to interact with task performance, specifically task performance in teaching arithmetic, was number anxiety or number avoidance. There was no direct evidence from previous studies that this characteristic was involved in performance, but the fact that intermediate teachers increased in arithmetic performance under supervision, together with evidence from Dutton's (6) early studies showing that elementary teachers have avoidance attitudes toward mathematics, suggested that the effect of supervisory activity might be either to reduce number anxiety or else to increase the magnitude of the approach motivation so that it became superordinate to the avoidance motivation, thereby bringing about a change in motivation ultimately leading to a performance change among intermediate teachers. Unfortunately, the fact that supervision also influenced reading performance was not given adequate weight when the hypothesis concerning number anxiety was formed. Clearly a reduction in number anxiety would not be expected to influence performance in teaching reading, but performance in reading did in fact change.

A second type of consideration entering instrument development, specifically task development, stemmed from the fact that tasks in only reading and arithmetic had been used. While these two skill subjects are of central importance in the elementary school, it is also true that such content areas as science and social studies are significant features of the elementary curriculum. The development of tasks in these latter two areas was therefore projected. The development of problem-solving tasks in social studies was never fully undertaken, however. Investigation of the curricular content of social studies in the primary grades among practicing teachers indicated that this content is so diverse that there was practically no hope for constructing a set of tasks for which any claim of representativeness could be made. The development of problems solving tasks in science presented a similar, but less severe problem, and ultimately full efforts were devoted to attaining unbiased science tasks, rather than splitting the efforts between science and social studies, since the probable pay-off in social studies seemed very low.

In sum, the situation at the outset of the study with respect to instrument development was as follows:

1. To develop a total battery of tests and scales which required not more than two hours and fifteen minutes for the average teacher to perform.

2. To develop a set of tasks in arithmetic, reading and science for intermediate teachers and a comparable set for primary teachers.

3. Develop an abstract (subject-matter free) scale or set of scales to assess concept attainment and sequencing or ordering skills, or a test of induction.

4. To develop a scale associated with teacher involvement or career commitment, and a scale associated with number anxiety.

5. To shorten and reconstitute the scales in the TCS.

Although the situation at the beginning of the project called for the revision and development of a substantial number of tasks and scales, it was also apparent that not all of the tasks and scales developed could ultimately be included in a battery of tasks and scales the total performance time of which was approximately 2.25 hours. Some type of strategy was therefore required not only for task and scale development, but for final selection. Excluding revision of the TCS, the strategy adopted was a cycle of task and scale proliferation--sub-study--revision--sub-study--scale reduction or elimination. For the TCS the strategy adopted was a recurring cycle of revision--pilot study--revision. In succeeding sections, the various tasks and scales developed are described first, then the various sub-studies pertinent to them outlined. Subsequently the revisions of the TCS and the principle sub-studies related to these revisions are described. The major pilot studies for the TCS, however, are described in Chapter 2 rather than in this appendix.

Development of the Problem-solving Tasks and Related Scales and Tests.

Arithmetic Tasks. Two tasks in teaching arithmetic, of known validity, were available from previous studies. Each of these tasks was not only a predictor of teaching success among intermediate teachers, but teachers scoring above the median on them were known to obtain better pupil gains in arithmetic than teachers scoring below the median on them (20). One of these tasks, hereafter labeled IA-3 (Intermediate Arithmetic 3), requires the teacher to examine a mixed exercise done by a pupil in grade 5, then rate each of ten statements about errors appearing in the exercise on a three point scale according to how much emphasis she would place on each type of error in a follow-up interview with the pupil. The processes involved in this task are finding the errors, inducing the category to which they belong, examining the frequency of errors to hits by the pupil in the category, and judging significance in accord with the frequency of hits to errors. The second of the tasks (IA-4) requires the teacher to rank seven division examples common to grade 5, but also found in grade 6, and to some degree in grade 4, according to their difficulty for pupils. Four concepts of what makes a division problem difficult for pupils (based on error rates) are required to correctly order the series.*

Because of the previous success of these tasks, the arithmetic tasks for primary teachers were modeled after them in general format.

---

*These problems appear in the traditional algorithm; examination of alternative algorithms now in use suggested that a change in algorithm would have modest effects on the task.

To develop the first of these tasks, PA-3 (Primary Arithmetic-3) a number of modern mathematics texts were examined, and a set of addition and subtraction examples involving three digit numbers, some with regrouping required in the units and tens, were selected. In the accelerated texts, problems of this type fall in the second semester of second grade, while in the "slower" texts they fall in the third grade. Subsequently, error rates were fixed for each category of example. After the exercise was thus set, a number of workbooks were searched for exercises which were in fact relevant to the errors, or else looked as if they might be relevant, but were not in fact. Ten such exercises were selected, some in traditional form, some in modern form (these frequently requiring diagrams and pictures of concrete materials). In the task, these exercises are rated by the respondent according to their relevance to correcting the errors made by the pupil. The processes involved in this task are highly similar to those required in task 3 for intermediate teachers, except that judging the relevance of exercises to categories of pupil errors rather than judging the relevance of statements about pupil errors to pupil errors is required. The former procedure was more time consuming, and more difficult than the latter, probably because the objective of each exercise must be inferred.

The second task for primary teachers (PA-4) was developed in parallel with the corresponding task for intermediate teachers. In it the respondent is required to rank eight addition problems at second and third grade level according to difficulty. The empirical error rates for addition problems of this kind were not available, and the "correct" sequence was determined by the standard sequences in texts and by concurrence of experts in arithmetic instruction.

Reading Tasks. Of the tasks in teaching reading available from previous development only one showed sufficient evidence on the criteria of score increases with the first two years of teaching experience and significant correlation with teaching success to be retained in the project. The task (PR-IR-1) requires the respondent to group and label ten errors made by a pupil while sight reading a list of 36 high frequency (Dolch) words. This task was retained for both primary and intermedicate teachers.

The second task (hereafter, "word ranking task") was also developed for both primary and intermediate teachers, and consisted of two sets of four words each.* The words in each set were drawn from the categories of words given by Gray (8) as words requiring increasingly sophisticated word attack skills on the part of the pupil in order to "unlock" them on first encounter. The problem for the respondent was to rank the words according to the level of difficulty they presented in the application of word attack skills. This task, one may note, was directly comparable to the arithmetic tasks requiring the ranking of arithmetic problems according to difficulty, and thus represented an ordering or sequencing type of task.

---

*Five words initially appeared in each set, but in several dry-runs, undergraduates were found to be unable to reliably rank five words on the criterion stated.

The third reading task was developed for primary and intermediate teachers separately, although identical procedures were used in each case. To develop these tasks (PR-2 and IR-2) a paragraph at a difficulty level of second semester second grade and a paragraph at second semester fifth grade were written, together with the appropriate comprehension questions.* These paragraphs were then taken to an elementary school, where the writer listened to 15 primary students and 15 intermediate students individually read aloud and answer the comprehension questions over the respective paragraphs. All word errors, part-word errors, and word omissions were recorded, as were the precise answers to the comprehension questions, thus providing a protocol for each student. These protocols were then analyzed, and four protocols at the primary level and four at the intermediate level chosen. The protocols represented the range from both systematic word errors and comprehension difficulties through no comprehension difficulties and few word errors.

Subsequently, a number of reading workbooks were searched to find examples of exercises actually relevant and seemingly, but not actually, relevant to the systematic word errors in the protocols. In addition, several paragraphs were drawn from readers, no longer in use, and slightly re-written to insure that they accurately reflected particular levels of reading difficulty. The exercises and the paragraphs were then placed with the protocols to make up the reading task. The problem for the respondent in this type of task is complex. The word errors must be classified and a decision made about the comprehension level of the student relative to the paragraph he read (i.e. the protocol), for each of four students. Next all exercises and paragraphs must be examined, and a decision made for each student concerning whether he should read an easier or a more difficult paragraph to check his comprehension, and whether he should do one or more exercises to check his work recognition or word attack skills. In essence, this type of task is a diagnostic task.

As the reader may indeed suspect, the early dry runs with these tasks among undergraduates indicated that each one consumed about forty minutes to resolve, and for undergraduates presented a problem of incredible difficulty, a part of it being able to hold the protocols of four students in mind while examining exercises and paragraphs for each. Interestingly, in these dry runs speech and hearing students, with considerable background in the diagnosis of speech difficulties, did substantially better than the typical undergraduates in elementary education. It was clear, in any event, that these tasks would not be functional as designed, and the two protocols yielding least discrimination in relation to the total score for all protocols were removed. In addition, the number of exercises and paragraphs were reduced.

---

*These paragraphs were quite carefully constructed, with the principal words drawn from the Spache (19) test for establishing reading level, and the reading difficulty level established by Dale-Chall formula.

3

Science Tasks. The development of tasks in the teaching of science was discovered to be quite difficult. The difficulty lay in constructing a tasks which did not explicitly require a specific type of knowledge, for example a knowledge of the simple machines, or of a set of biological relationships, or of the solar system. The source of this difficulty lies in introducing into the task a specific achievement component which, if absent for a respondent, precludes successful performance of the task. To avoid this difficulty, a single type of science task, which appeared to provide equal opportunity for all, was constructed for primary and intermediate teachers separately.

To construct the tasks, a list of some 6,000 questions posed by children in elementary schools was consulted. A list of 25 questions at the primary level and a list of 25 at the intermediate level were then selected such that sub-sets of questions surrounding a single concept, or a principle appropriate to the age level, could be formed. Not all of the questions could be placed in a particular sub-set, however. These questions functioned as distractors. Instructions to respondents in these tasks are first to suppose that it is near the beginning of the school year and that the class has raised the questions listed. Subsequently, the respondent is asked to group the questions in such a way that they undergird possible units in science which might be developed during the subsequent weeks of school, and to label the groups.

It is apparent that the response latitude in a task of this type is quite wide, and potentially presents great scoring difficulties. To minimize these difficulties, both the primary and the intermediate tasks were performed by a graduate elementary education class in research in teaching science and social studies. The resulting responses were then plotted in a 25 x 25 matrix for each task, primary and intermediate, so that the frequency with which a particular question was associated with any other question could be obtained. Because the questions had initially been selected around particular concepts and principles, the frequencies in the matrix produced clusters showing the questions associated with each principle or concept. The questions falling in these clusters were then counted as "hits." Some questions were tangentially related to certain principles or concepts, however, and received moderate frequencies in relation to particular clusters.

These became "neutral" responses, while questions irrelevant to a cluster were recorded as errors. A scoring scheme of hits minus errors with the "neutrals" not counting was then established for these tasks.

In the dry runs done with the tasks, several interesting patterns of responding were observed, each appearing to be linked to a particular cognitive style. One pattern, labeled the "mediating" pattern was relatively common among untutored undergraduate students. This pattern consisted of placing the "hits" in a unit, then vastly extending the number of questions to be embraced in the unit in a fashion which distinctly suggested that a particular word in a question functioned as mediator to join that question to another. For example, the question "How do fish live under water"? was associated with "What makes things get rusty"? and with "How can clouds carry rain"? which was associated with "Where doclouds come from"? and so on and on.

A second pattern was approximately the reverse of the first, namely, many very small or restricted units were formed. Respondents with this approach might be labeled over-analytic or "fractionating." The third, and predominant, pattern was that of constructing units containing intermediate numbers of questions. It is units of this length which form the foundations of the clusters found in the analysis of the matrix.

A number of different scoring systems were created in an attempt to take into account quite different approaches to the science tasks. The systems were quite cumbersome, however, and were ultimately abandoned in favor of the simpler hits-errors scheme. It should be noted, nonetheless, that scoring the science tasks by this scheme tended to penalize both the "mediating" and "fractionating" respondent without substantial evidence that such respondents might be less proficient in teaching science in the elementary school.

Induction-Related Tasks and Scales. Four types of instruments, one of which was borrowed and three of which were created, were employed in the early phases of the project. The first, and borrowed, instrument was the Thurstone Letter Sets, of which two parallel forms of 15 items each were used. The Letter Sets may perhaps be best regarded as measures of concept attainment. The second instrument, which appears in Appendix 2, was constructed by the writer in an effort to reveal the number of different sequences or orders the respondent could produce from a single sequence. A letter sequence, a letter sets sequence, a word sequence, a numerical sequence and a pictorial sequence were used. While the original intent of this task was to determine the scope of the sequencing behavior of the respondent, it soom became apparent that a distinct "creativity" element was built into instrument in that a kind of "sequencing fluency" was required in order for the respondent to generate score. The feature of the instrument created conceptual difficulties in that creativity and sequencing elements were confounded, with little hope of extricating each short of an extended factor analysic study. A second feature of the instrument that created difficulty lay in scoring. While each scorable sequence had to be generated by a rule, the number of sequences generated by respondents in the dry runs led to an interminable amount of rule-discovering by the scorer since all possible rule-governed sequences had not been ascertained beforehand.

The third instrument developed was a number series induction test. Two forms of this test were developed, one in which the respondent had the option of indicating that a series had no rule, and some series did not have a rule (other than randomness), and one in which all series had a rule and no alternatives were given to finding the rule.

The fourth instrument in this group was developed after the others were abandoned, and represented to some degree a combination of the ideas underlying the earlier devices. This test was based on a core of verbal induction items to which several number series items and four "alternate rule" items were added. The "alternate rule" items were standard verbal induction items constructed in such a way that two rule-governed sets could be formed from them rather than the single set customary in items of this type. A sample appears below.

148. Which of the following seems to belong least to the group?

148-1. Elephant 148-2. Horse 148-3. Camel 148-4. Tiger
148-5. Llama

149. Using a different concept of principle from the one used in the above item, which word seems to belong least to the group?

149-1. Elephant 149-2. Horse 149-3. Camel 149-4. Tiger
149-5. Llama

Number Anxiety. Work on the number anxiety scale began on the basis of an item extracted by Dreger and Aitken (5) in a factor analysis involving the Taylor Anxiety Scale, namely, "Many times when I see a math problem, I just freeze up." A number of items were then written to extend the scale, and after an initial dry run all were buried in a set of items from the Alpert and Haber (1). . Facilitating versus Debilitating Test Anxiety scale, and items from an experimental audience anxiety scale. Following a sub-study, the number anxiety scale was abandoned, for reasons to be described at a later point, and a "mathematics attitude scale" developed.

Career Involvement. As noted on an earlier page, an initial career involvement scale was developed by White and used in conjunction with beginning teachers. This scale met several validity criteria, including associations with changes in task performance, ability to distinguish teachers chosen by supervisors as "highly committed to teaching" from those chosen as "not committed to teaching," and ability to differentiate married women who continued to teach from those who dropped out of teaching. There were, however, two difficulties with the scale. The first was that it was constructed in such a way that one could not logically respond to some items unless he had had teaching experience. The second difficulty was the apparent openess of the items to faking. These difficulties were thus in need of correction. A second scale developed by White in the same study involved attitudes toward marriage versus career. This scale produced reasonably interesting results among women undergraduate in elementary education, revealing, for example, that women of middle and upper middle class origin were more marriage and less career oriented, while women of lower middle and working class origin were more career and less marriage oriented, but its use with experienced elementary teachers failed to yeild validity on the same criteria against which the career involvement scale was initially validated. Some interest was nonetheless retained in the scale, partly because there appeared to be a distinct tendency in an earlier study for school systems in upper middle class communities to employ beginning female teachers of middle class origin, who then dropped out after the first year of teaching.

During the first year of the project, work on the two scales noted above was divided, with development of the career involvement scale retained in the project, while the development of the marriage versus career scale was continued by Mills (13) as part of a doctoral dissertation. In revising the career involvement scale, two parallel sets of 14 items each were written, one set worded for undergraduates wihout teaching experience and the other set worded for persons with teaching experience. The set for undergraduates were then subjected to a dry run and the results used for item analysis on the criterion of association with total score. The results of this dry run were of particular interest, not only because most of the items were homogeneous and discriminating, but because a particular form of item seemed to obtain the best results, namely, the pairing of an item using the first person pronoun in a statement of commitment against an item which in essence provided an escape route from commitment. Two examples appear below.

4. ______ a. I find education courses interesting.

______ b. Teaching is a good profession but I would not like to teach all the rest of my life.

8. ______ a. As a teacher I would expect to regularly devote an hour or so to pupil assistance outside of class hours.

______ b. Administrators such as principals are frequently chosen more for their seniority than for their suitability for the position.

Sub-studies Using the Tasks and Scales Developed. In the project, a "sub-study" lay between a try-out or a dry run and a full scale pilot study. The central difference between a sub-study and a pilot study may be viewed as one of design. In a sub-study, the design was adapted to the samples of subjects available for study while in the pilot studies, for the most part, a general design was created then the relevant populations sought.

Sub-study 1. The first sub-study was conducted near the end of the first semester 1964-65, with approximately 70 undergraduates all of whom were completing an arithmetic methods course, and approximately 25 of whom were also enrolled in one section of methods course in language arts. The students in arithmetic methods course were administered the total set of arithmetic tasks for both primary and intermediate teachers, the Thurstone Letter Sets and the tests of ability to generate multiple sequences from a single sequence. In addition, those in the reading class were administered the word-ranking task in reading.

Frequency distributions for each test and scatter-diagrams between the Letter Sets, the sequencing test and all arithmetic tasks and the reading task were then plotted, and correlations were computed when a scatter-diagram indicated that a significant correlation might appear. While the distributions of the variables involved was approximately normal, no association appeared between the sequencing test (or any part thereof) and any of the arithmetic tasks or the reading task. A correlation of .41 appeared between the Letter Sets and sequencing test, however, and, in addition a small positive association between the Letter Sets and the diagnostic tasks in arithmetic. Finally, the correlation between the two forms of the Letter Sets was .55. The situation thus appeared to be that such true-score variance as was available in the sequencing test was associated with the diagnostic arithmetic tasks. The sequencing test was then abandoned since there was no way to gain a foothold for item analysis, other than to the Letter Sets, which would clearly have created a redundancy.

Because the performance time of the tasks and Letter Sets was rather long, one form of the Letter Sets was deleted and the diagnostic arithmetic task for primary teachers (PA-3) was shortened and the format changed.

Sub-study 2. This sub-study was conducted during the spring semester, 1964-65, and consisted of three parts, one associated primarily with arithmetic, one associated with reading, and one associated with the TCS, the latter part will be described at a later point.

The arithmetic aspect of this study involved all arithmetic tasks for both primary and intermediate teachers, the number anxiety scale, the test anxiety scale, all administered on a pre-test, post-test basis, and the Letter Sets, administered as a pre-test. Eighty students taking arithmetic methods served as the subjects. The results may be observed in Tables 1 and 2.

The central analysis in this study was a 2 x 2 x 2 repeated measure analysis of variance with the upper and lower 25 percent of the students on number anxiety serving as one dichotomy, a median split on facilitating test anxiety serving as the second dichotomy, and pre-test, post-test scores as the third dichotomy.

The results in Table 2 indicate that both number and test anxiety were consistantly enough associated with task performance to produce significant effects across both pre and post tests, although these effects were not strong. The disturbing aspects of this analysis lay, however, in the magnitude of the within subjects error term (error 2) which was quite substantially greater than the between subjects error term (error 1). This state of affairs suggested that an unidentified variable was probably operating in the change scores. This variable might have been a change in the number anxiety scores from pre to post test, but an analysis of these changes revealed that means were approximately identical from the pre to the post tests. Moreover, re-examination of the items in the number anxiety scale suggested why no changes had occured with the methods course, namely many of the items dealt with the personal history of the respondent and thus made a change of response virtually impossible. To correct this problem, a decision was then made to change the scale from a number anxiety to an attitude toward mathematics scale which would have in it the possibility of change.

Analysis of other aspects of the data indicated that the Letter Sets did not correlate significantly with task performance, and the decision was made to change to numerical induction tasks to accompany the arithmetic tasks.

The aspect of this sub-study involving the reading tasks was originally intended to produce relationships between the reading and the arithmetic tasks. However, the diagnostic reading tasks were employed for the first time in the study, and it was at this point that the tasks, as orginally constructed, were discovered to be much too difficult, with students responding approximately at random to two of the four pupil protocols used. These two protocols were then deleted and other modifications made to shorten the response time to the two remaining protocols.

Table 1. Mean Performance on Arithmetic Teaching Tasks as a Function of Number Anxiety and Test Anxiety.

Facil. Test Anx.	Number Anxiety High Pre	Number Anxiety High Post	Number Anxiety Low Pre	Number Anxiety Low Post
High	19.40	22.40	23.00	19.10
Low	14.10	17.90	18.20	23.20

Table 2. Repeated Measures ANOVA for Table 1.

Source	SS	df	MS	F	p
between S's Number anxiety	117.62	1	117.62	4.10	05
Test anxiety	137.82	1	137.82	4.81	05
No. anxiety by test anxiety	103.50	1	103.50	3.61	ns
Error 1	1031.95	36	28.66		
within S's					
between tests	78.02	1	78.02	2.00	ns
No. anxiety by tests	40.60	1	40.60	1.03	ns
Test anxiety by tests	117.60	1	117.60	3.01	ns
No. anxiety by tests anxiety by tests	8.79	1	81.79	2.09	ns
Error 2	1405.99	36	39 06		

Following this sub-study, the number anxiety scale was reconstructed into a mathematics attitude scale and tried out on a group of elementary teachers enrolled in graduate course and arithmetic methods along with the arithmetic tasks. The correlation between the scale and the tasks was only .03, however, and additional revision was then undertaken.

Sub-study 3. This sub-study was conducted in the fall of 1965 across multiple sections of three methods courses: arithmetic, science and language arts. The students in the arithmetic methods course were pre and post tested on all arithmetic tasks and the revised mathematics attitude scale. In addition, the career involvement scale, and a preliminary form of the marriage versus career scale was administered during the middle of the course. Students in the science course were pre and post tested on the science tasks, while students in the reading course were tested over the reading tasks near the end of the course.

In addition, Dr. William Powell at the University of Illinois used the reading tasks on a pre-post test basis with two section of students in a reading methods course at that institution.

The results of the study of arithmetic course indicated, first, that there was no association between the arithmetic attitude score and performance on the arithmetic tasks. Slight changes did occur on the arithmetic attitude score, but these changes were not significant, and there was no indication that attitudes toward arithmetic were significantly associated with changes in arithmetic task performance. Second, no association could be found between the numerical induction test and arithmetic task performance. These findings were largely consistent with previous findings and suggested that both numerical induction and arithmetic attitudes were, at best, weak influences on task performance.

Examination of arithmetic task performance in relation to career involvement indicated, however, that career involvement had a distinct influence on arithmetic task performance. First, it was significantly and strongly ($p < .01$) associated with pre-test performance on the arithmetic tasks. Second, changes in performance during the course were wholly associated with career motivation, with the group below the median in this variable increasing approximately two-thirds of a standard deviation in performance between the pre and the post test.

Because approximately half of the students enrolled in arithmetic methods were also enrolled in the science methods course, the influence of career involvement could also be examined with respect to the science tasks. The results were almost identical to those for arithmetic, with career involvement being associated with pre-test performance on the science tasks at $p < .01$ and the below the median group in career motivation gaining about two-thirds of a standard deviation on the science tasks with the course.

Subsequently, the effects of career motivation on the pre-tests in science and arithmetic were examined with the primary and the intermediate teachers, and tasks, separated. The pattern of results indicated that the effects tended to be grade and task specific, that is the greatest influences of career motivation for primary teachers appeared in their performance of the primary tasks while the greatest effect for intermediate teachers reappeared in their performance of the intermediate tasks. This result suggested why a distinction between tasks for primary teachers and tasks for intermediate teachers should be maintained, namely, each group seems best motivated to perform those tasks specific to the grade range in which they are involved.

Other effects of interest from this sub-study were first, that career motivation was not associated with performance of the reading tasks; second, that numerical induction was not associated with performance of these tasks, and third, that career motivation was not associated with attitudes toward marriage versus career, and the latter was not associated with task performance.

In addition, the study conducted by Powell at the University of Illinois indicated that there were significant changes in the ability of students in a diagnostically oriented course in teaching reading to perform the diagnostic reading tasks, but not in the word-ranking or sequencing task in reading. These results suggested, but did not conclusively establish, that a problem existed with the work-ranking task. Intensive examination of the diagnostic tasks and the scoring system, by Powell, also revealed a number of shortcomings in them, and changes were subsequently made in task format and scoring to correct these shortcomings.

Between the third and fourth sub-studies, the career involvement scale was item analized and revised, a verbal induction scale was created and dry-run, the numerical induction scale was item analyzed and extended. Following these revisions, a battery for primary teachers and a battery for intermediate teachers were formed such that each group had science, arithmetic and reading tasks appropriate to its grade range, while all other scales, namely, verbal induction, numerical induction and career involvement were common across grade ranges.

Sub-study 4. The purpose of the sub-study was to ascertain the time requirements for performing the various instruments, and to examine the interrelationships among them. The study conducted among undergraduates completing reading, arithmetic, science and methods courses who volunteered to take the battery appropriate to their grade level preference on one of two evenings. The volunteers were then signed, and the testing arranged; unfortunately, however, many of those volunteering did not appear for testing, so that an N of only 35, intermediate and primary teachers pooled, were actually tested. This factor badly damaged the study since separate analyses of the data for primary and intermediate teachers were precluded, and a correlational analysis of the scores of the two groups pooled was accomplished only through Z score transformations of the major scores for each group.

The results of this study did, however, have certain significant consequences. First, the mean performance time was in excess of two hours, approximately one hour longer than the maximum performance allowable in the field for the direct (face to face) testing of teachers. Thus a drastic reduction in tasks and scales was imperative. Second, the correlations among the tasks and scales indicated that the numerical induction scale was not significantly associated with the Z score for the arithmetic tasks, but the latter was associated with the Z score for the science task and the diagnostic reading task. The work-ranking task in reading, however, showed negative associations with the diagnostic reading task, and the science and arithmetic Z scores. A distincly disturbing state of affairs.

As a result of this study, the numerical induction test was dropped and the verbal induction test revised and moved into the TCS as a power test. In addition, the format of the diagnostic tasks in reading and arithmetic was modified to reduce response time.

It should be noted that while sub-study 4 was underway, a pilot study, conducted by Rex Brown, was initiated with a full battery of tasks in teaching reading, and that the final revision of the reading tasks issued from the results of this study, which is described in Chapter II of the report.

Revision of the Teacher Characteristics Schedule. As originally constructed by Ryans', the Teacher Characteristics Schedule contained 350 items to which approximately 650 responses were required since some items required three responses. From these items 10 scales were extracted:

Xco	friendly, understanding vs cool, aloof classroom behavior;
Yco	responsible, business like vs slipshod classroom behavior;
Zco	stimulating imaginative vs dull, routing classroom behavior;
B	permissive, child-centered vs traditional, subject-centered educational viewpoint;
R	attitude toward pupils
Rl	attitude toward democratic pupil practices
Q	attitude toward school staff
I	verbal intelligence
V	validity of response
S	emotional stability

The extraction of the scales from the items was accomplished by a key for each scale (score key 111) and each key was constructed in such a way that responses to individual items were scored on more than one scale. That is, a single response might be counted on as many as three scales. The overlap among the scales was, of course, very high, invariably producing significant correlations among all the scales among elementary teachers, and in the study of beginning teachers immediately preceding the present study, invariably producing a single factor.

Because of the original TCS was long and time consuming and relatively undifferentiated for elementary teachers, the objectives for revising it were both to decrease its length and increase the independence of the scales while maintaining reasonable reliability. These objectives were accomplished in a series of studies extending over a three year period, of which only the first three sub-studies are reported in this Appendix, while the major pilot studies are reported in Chapter 2.

Sub-study 1. In the first phase of this sub-study, four moves were made. First, scales I and V were dropped from the TCS. I was dropped primarily because it loaded with the personal-social scales in the TCS rather than with arithmetic and reading task performance in the preceding study of beginning teachers. V was dropped because no evidence of its validity have ever been produced. Second, all factual items in the TCS were eliminated. These items covered such matters as the grade level in which the respondent taught, the kind of community in which he taught, and so on. Third, the remaining items were then item-analyzed, using the responses of the beginning teachers from the preceding project, on the criterion of association with the total score. This item analysis revealed that a large group of pictorial items used as quasi-projective items in the original TCS were not associated with the total scores of the scales to which they were ordered, and all such items were eliminated. In addition, the items on scales R and R1 were found to overlap so greatly that these two scales were collapsed into a single scale. Finally, all items on any scale which failed to associate with the total score for that scale at the 20 percent level or less under Chi Square, using the upper and lower 27 percent of those scores on the scale as the criterion groups, were eliminated. Fourth, the items of Scales X and Y, which predicted the presence of problems with discipline and in teaching reading among beginning teachers, were further item-analyzed by utilizing the group with discipline problems and reading problems as one criterion and the group with no problems as beginning teachers as the other criterion. Items which discriminated between these two groups at the 20 percent level or less by Chi Square were then added back into the appropriate one of these two scales if they had initially been deleted for failing to associate with the total scores on their scale. The latter move, it should be noted, potentially enhanced the future validity of each scale, but also increased the heterogeneity of each.

The moves described above produced a residual group of 140 items distributed over seven scales; the total scores points over the seven scales was 336, however, indicating that very great overlap among the scales continued to exist. To put the matter another way, most of the discriminating items discriminated on more than one scale, suggesting that a large source of common variance existed among sub-sets of scales if not among all scales.

In the second phase of this sub-study, a set of approximately 90 score protocols were drawn from a study by Reisert (15) in which the original TCS had been used with beginning elementary teachers. These protocols were then re-scored using the newly derived scoring keys, and the split-half reliabilities of each scale calculated. The results seemed disastrous, with most of the reliabilities running less than .50, and the reliability of the revised Y scale being negative (if one can call a negative _r_ a reliability). Reliabilities of magnitude obtained clearly indicated very great item heterogeneity, but a part of the problem was interpreted to lie in the nature of Reisert's sample, which was one-third regularly certified teachers and two-thirds non-certified teachers, an extraordinarily heterogeneous group, a substantial part of which was drawn from secondary teachers who had switched to elementary following graduation.

During the interval in which the first sub-study was being conducted, a number of additional items which were ultimately to be placed in the B scale (child-centered vs subject-centered) were under construction. These items were initially placed in a scale named "authority-sharing vs authority-centering" which was tried out in a group of undergraduates at Miami University by Dr. M.E. Fakouri. Subsequently, 10 items from this scale were selected as holding some promise, and placed with the 140 items from the TCS, resulting in a new form (E65) with 150 items.

Sub-study 2. Like the foregoing sub-study, this one had two phases. In the first phase 115 undergraduates in elementary education at the junior level were administered Form E65 of the TCS. Each item was then correlated (bi-serial) with the total score for its scale, and in addition, the percent correctly responding to each item was obtained. The data resulting from these analysis were reasonably complex since virtually every item of the 140 drawn from the original TCS was ordered to more than one scale. The problem thus became to eliminate an item from a scale with which it did not correlate and retain it in a scale with which it did correlate. This problem was resolved by ranking all items in each scale according to their bi-serial _r_ and eliminating from the scale those items which correlated less than .20 with the total score for that scale. Under this procedure any item which failed to yeild a bi-serial _r_ of .20 on any scale was eliminated from the schedule, but any item which correlated with any scale .20 or more was kept in that scale, thus, some (in fact, many) items were retained in more than one scale since they were originally scored by Ryans' in that scale and continued to correlate with the total score for the scale.

After the items had been ordered to scales, items within scales were matched on the basis of bis-r and percent passing so that two equal halves were hypothetically produced.

In addition to the above procedures, the items remaining after item analysis were also subjected to intercorrelations by means of phi coefficients, and the resulting matrix of phi coefficients favored by the principal components method with oblimin rotation to oblique factors. The results were, however, completely uninterpretable, and the notion of extracting a new set of homogeneous scales from the items was set aside.

In the second phase of this sub-study, Form E65 of the TCS was administered to 48 graduate students enrolled in a summer school course. The resulting protocols were then scored on the split score keys developed from the analysis of the undergraduate sample. The results were as follows.

Scale	No. of Items*	rxx
X	28	.66
Y	28	.56
Z	24	.67
R-R1	34	.76
Q	30	.77
B	30	.60
S	30	.71

During the period that the reliability studies of Form E 65 were being conducted, work was also continued on the authority-centering vs authority-sharing scale, for which a number of new items had to be written. At the same time, the career involvement scales were under development as described in the earlier sections of this Appendix. Following the reliability study described above, both the career motivation scale for teachers (14 items) and the authority-centering vs authority-sharing scales (16 items) were ordered to the Schedule, which had been by this time reduced to 120 items from the original TCS. The resulting form of the revised TCS was called Form E66, and with it a number of pilot studies were initiated, as described in Chapter 2.

---

*This is the number items scored. Since an item could be scored on more than one scale, the sum of the scored items across scales is greater than the number items in the Schedule.

APPENDIX 2

Data Sheet for Teachers

1. How many years have you taught in your present school? ______. In your present school system? ______

2. Have you taught in other school systems? yes no. If "yes," how many years? ______. What was the last year you taught in another system? ______.

3. Do you hold the B.A. or B.S.? yes no; year completed ______. The M.A. or M.S.? yes no; year completed ______. The M.A. or M.S. plus 30 hours of credit beyond? yes no; year completed ______.

4. Which grade(s) have you taught other than the one you now teach?

5. With respect to your teaching, in which school subjects do you feel you have greatest interest and teaching skill?

   In which school subjects, if any, do you feel you have least interest and least teaching skill?

6. Did you participate in an inservice program sponsored by your present school system during the past 18 months? yes no. If "yes;" which programs were most valuable to you?

TEACHER FORMAT - DATA BDT-E'67

Name ____________________
Home Address ____________________
____

Item		Column	Punch	Item		Column	Punch
1.	Card number	1	____	19.	Sex	23	____
2.	System ID	2-3	____	20.	S-E of school	24	____
3.	School ID	4-5	____	21.	Task 1 (raw score)	25-26	____
4.	SKIP	6	SKIP	22.	Task 2 (raw score)	27-28	____
5.	Teacher ID	7-9	____	23.	Reading total (Z score)	29-31	____
6.	Grade	10	____	24.	Task 3 (raw score)	32-33	____
7.	Rating	11	____	25.	Task 4 (raw score)	34-35	____
8.	Problem/model	12	____	26.	Arithmetic total (Z score)	36-38	____
9.	Years in present school	13	____	27.	Science hit (raw)	39-40	____
10.	Years in present system	14	____	28.	Science error (raw)	41-42	____
11.	Other system experience	15	____	29.	Science H-E+20 (Z score)	43-45	____
12.	Total experience	16	____	30.	Total task (Z score)	46-48	____
13.	Degree	17	____	31.	Resourcefulness (Z)	49-50	____
14.	Taught other grades	18	____	32.	Organization (Z)	51-52	____
15.	Reading strength	19	____	33.	Viewpoint (Z)	53-54	____
16.	Arithmetic strength	20	____	34.	Stability (Z)	55-56	____
17.	Science strength	21	____	35.	Involvement (Z)	57-58	____
18.	Inservice experience	22	____	36.	Induction (Z)	59-60	____

Interview Schedule for Principals

Questions:	Responses:
I. BACKGROUND OF PRINCIPAL	
A. Years as principal in this school?	Years (excluding this year) _____
B. Other capacities in this system?	In this school In other schools
C. Other experience and capacities including grades taught (in other systems).	OTHER EXPERIENCE
D. Institution and year of M.S. Degree	Institution __________ Year _____
E. Number of graduate hours beyond M.S.	Graduate Hours ____________
F. Sex	Male _____ Female _____
II. SOCIAL SCHOOL SITUATION	
A. Number of pupils grades 1-6	Number of pupils: ___________
B. Occupational background of parents. which of the following categories are most numerous in the student body?	OCCUPATIONAL STATUS OF PARENTS
1. Upper middle and above Professional managerial (medium and large business)	1. _____
2. General middle teachers, small business, sales (employed), clerical	2. _____
3. Farm owners and manager	3. _____
4. Skilled labor	4. _____
5. Semi-skilled & unskilled	5. _____
6. Unskilled, sporadic employment, on relief	6. _____ Mixed (specify)
C. Are the parents in your school easy or difficult to work with?	Easy _____ Difficult ______ Comments:

	What kinds of difficulties?	Difficulties:
	What kinds of parents are easy to work with?	Type of parent:

III. PROFESSIONAL DUTIES:

A.	When serious problems arise in your duties as principal, what official do you confer most in order to work out a solution?	Office: Name: Title:
B.	Is there anyone who does direct classroom supervision besides yourself:	Yes ______ No ______
	What is his title?	Title:
	Is this a system-wide position?	Yes ______ No ______
C.	Approximately how much time does ___(official)___ spend on the average per week in direct classroom supervision?	Average Number of Hours ________
D.	What methods does ___(official)___ most often employ? (e.g. conferences, visitation, etc.)	Methods:
	Does ___(the official)___ usually report his progress in supervising?	Yes ______ No ______
	What type of reporting does he usually make?	Type of report:
	To whom does he report?	Official's title:
E.	Think of some of the outstanding teacher difficulty cases that you have had to deal with in this school. What kinds of difficulties did these teachers have?	Teacher difficulty:
F.	Approximately how many hours per week, on the average, do you spend in direct classroom supervision?	Hours classroom observation:

G.	What methods of supervision do you most often employ?	Supervising methods:
H.	While supervising, what sort of things do you look for to determine whether a teacher is doing a good job or not?	Checklist:
	Are these characteristics a part of some type of checklist which you use during supervision?	Yes ______ No ______ Explain "yes":
I.	Thinking in terms of your present staff, which teachers have needed the most supervision this year? (Note: jot down names, then go through list, identifying problem.)	Teacher Problems Name Type of Problem
	About how many hours per week, on the average, have been spent supervising these teachers?	Hours Supervision: ____________
	Are there any teachers on your present staff that you would ask to help new staff members with teaching? What qualities do they possess that would cause you to choose them? (categories) Get names--qualities.	Teachers Models Name:

IV. OTHER RESPONSIBILITIES

A.	In the line and staff organization of this school system who is your immediate superior (line)?	Line Position Office: Name:
B.	Have your teachers engaged in in-service training programs?	Inservice Yes _______ No _________
	What kinds of in-service activities? (also when)	Activity: Date:
	Which of these activities are you required to conduct?	Required Activities:
C.	How frequently and when do you hold faculty meetings?	Faculty Meetings Frequency: ______________ When: ______________
D.	How much time do you spend with parents on the average per week?	Parent Contact Hours: ________
E.	Do you perform any guidance functions? Hours per week?	Guidance Yes ______ No ______ Hours _____
F.	How much independence are you given in running your school? (attempt to rate: None to complete)	Independence
G.	To what teacher organizations do your teachers belong? Have any difficulties arisen with any of these organizations? Which ones? What kinds of difficulties?	Organizations: Yes ______ No ______ Organization names: Difficulties:

H.	Do you have any difficulties with "teacher cliques?"	Clique Problems
		Yes ______ No ______
	Type of difficulties:	Types:
I.	What is the teacher turnover rete in your school?	Turnover Rate
		School rate: ________/ ________ leave total
	Rate in the system:	System rate: ________/ ________ leave total
J.	What is the ratio of the usual number of job applications to the number of job vacancies?	Application/Vacancies
		Ratio:
K.	When you interview a prospective teacher for a position in your school, what selection criteria do you rely on the most?	Teacher Personnel Selection criteria: (positive)
	What types of teachers do you try to avoid?	Selection criteria: (negative)

TEACHER BEHAVIOR WEIGHTING SCALE

Shown below are a number of statements of teacher characteristics, including skills, abilities and behaviors. While each statement probably expresses a generally desirable characteristic of teachers, the weight given to a particular characteristic as a factor in teacher success is likely to vary from school to school and system to system depending on the particular setting and situation. In order to find out what the weight of each characteristic is in your particular school setting, we must ask you to weight each characteristic relative to several others. We have therefore placed each characteristic in several different pairs. The characteristics in each pair are not opposites; rather, they are simply different.

To show which of the two characteristics in each pair has the greater weight as a factor in teacher success in your particular situation, proceed as follows: 1. examine each member of the pair; 2. choose that member of the pair which has the greater weight as a factor in teacher success in your situation; 3. show the amount of weight the chosen member has over the other member by circling "1", "2", or "3" on the scale next to the chosen member. Do not mark on the scale next to the other member.

Circling "1" means that you give only slightly more weight to the chosen member than to the other member.

Circling "2" means that you give considerably more weight to the chosen member than to the other member.

Circling "3" means that you give much more weight to the chosen member than to the other member.

PAIRS

Is skillful in adapting learning tasks to individual differences	3	2	1	1	2	3	Readily shows affection for pupils
Seeks widespread pupil participation in selecting and planning of units of work	3	2	1	1	2	3	Is skillful in diagnosing learning difficulties among pupils
Plans and organizes daily classroom program in advance	3	2	1	1	2	3	Seeks widespread pupil participation in selecting and planning of units of work
Maintains atmosphere of acceptance and understanding of pupils	3	2	1	1	2	3	Is skillful in arranging learning tasks in a carefully sequenced, logical order
Seeks to understand and reduce social-emotional problems of pupils	3	2	1	1	2	3	Establishes routines and limits that are well understood by pupils
Is skillful in adapting learning tasks to individual differences	3	2	1	1	2	3	Seeks to understand and reduce social-emotional problems of pupils
Is consistent and firm in managing pupils	3	2	1	1	2	3	Seeks widespread pupil participation in selecting and planning of units of work
Is skillful in arranging learning tasks in a carefull sequenced, logical order	3	2	1	1	2	3	Encourages and supports expression of ideas by pupils
Readily shows affection for pupils	3	2	1	1	2	3	Is consistent and firm in managing pupils
Plans and organizes daily classroom program in advance	3	2	1	1	2	3	Readily shows affection for pupils
Encourages pupils to form small working groups	3	2	1	1	2	3	Is consistent and firm in managing pupils
Encourages and supports expression of ideas by pupils	3	2	1	1	2	3	Is skillful in adapting learning tasks to individual differences

PAIRS

Seeks widespread pupil participation in selecting and planning of units of work	3 2 1	1 2 3	Is skillful in arranging learning tasks in a carefully sequenced, logical order
Readily shows affection for pupils	3 2 1	1 2 3	Is skillful in arranging learning tasks in a carefully sequenced, logical order
Is consistent and firm in managing pupils	3 2 1	1 2 3	Seeks to understand and reduce social-emotional problems of pupils
Encourages and supports expression of ideas by pupils	3 2 1	1 2 3	Establishes routines and limits that are well understood by pupils
Establishes routines and limits that are well understood by pupils	3 2 1	1 2 3	Maintains atmosphere of acceptance and understanding of pupils
Encourages and supports expression of ideas by pupils	3 2 1	1 2 3	Plans and organizes daily classroom program in advance
Establishes routines and limits that are well understood by pupils	3 2 1	1 2 3	Encourages pupils to form small working groups
Maintains atmosphere of acceptance and understanding of pupils	3 2 1	1 2 3	Is skillful in diagnosing learning difficulties among pupils
Plans and organizes daily classroom program in advance	3 2 1	1 2 3	Maintains atmosphere of acceptance and understanding of pupils
Is skillful in diagnosing learning difficulties among pupils	3 2 1	1 2 3	Encourages pupils to form small working groups

ORDERING TASKS

In this task your problem is to bring some kind of explainable, or meaningful order or organization to sets whose elements are disorderly or disorganized. In all of the sets shown below, it is possible to achieve more than one kind of explainable or meaningful order. Since you may not see all of the orders possible the first time you try a set, work through all of the sets, putting down the meaningful or explainable orders you see, then go back and work on the sets that you might be able to find more order for with additional effort.

SET 1 Arrange the letters in as many explainable, meaningful orders as you can.

R D E T I

Order 1 ____ ____ ____ ____ ____

Order 2 ____ ____ ____ ____ ____

Order 3 ____ ____ ____ ____ ____

Order 4 ____ ____ ____ ____ ____

SET 2 Arrange the words in as many meaningful orders as you can.

Seldom Ships Steady Seas Strong Sink

Order 1 ______ ______ ______ ______ ______ ______

Order 2 ______ ______ ______ ______ ______ ______

Order 3 ______ ______ ______ ______ ______ ______

Order 4 ______ ______ ______ ______ ______ ______

Order 5 ______ ______ ______ ______ ______ ______

SET 3 Arrance the following letter groups in as many explainable orders as you can. Each group has a number. To save time, use the number of the group in making your orders rather than copying each letter group.

RRROR	NNANN	XXXXU	EMMMM	LILLL
1	2	3	4	5

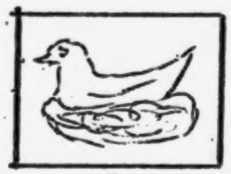
Order 1 ____ ____ ____ ____ ____

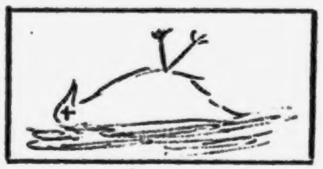
Order 2 ____ ____ ____ ____ ____

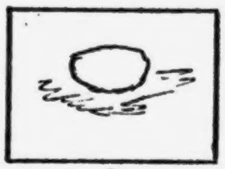
Order 3 ____ ____ ____ ____ ____

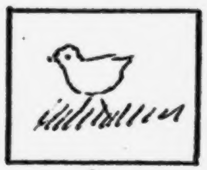
Order 4 ____ ____ ____ ____ ____

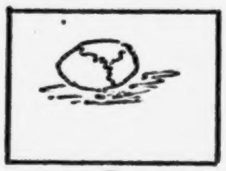
SET 4 Arrange the following pictures in as many explainable, meaningful orders as you can. Each picture has a number. To save time, use the number of each picture in making your orders.


1


2


3


4


5

Order 1 _____ _____ _____ _____ _____

Order 2 _____ _____ _____ _____ _____

Order 3 _____ _____ _____ _____ _____

Order 4 _____ _____ _____ _____ _____

Order 5 _____ _____ _____ _____ _____

Order 6 _____ _____ _____ _____ _____

SET 5 Arrange the following numbers and operators (+, -, x, =) in as many true statements or true equations as you can. Do not put down equivalent equations. One example is given.

1 3 5 7 9 x - + =

Equation 1 9 x 1 + 3 - 5 = 7 (equivalent, 3 - 5 + 1 x 9 = 7, etc.)

Equation 2 ____________________

Equation 3 ____________________

Equation 4 ____________________

Equation 5 ____________________

Equation 6 ____________________

Equation 7 ____________________

Equation 8 ____________________